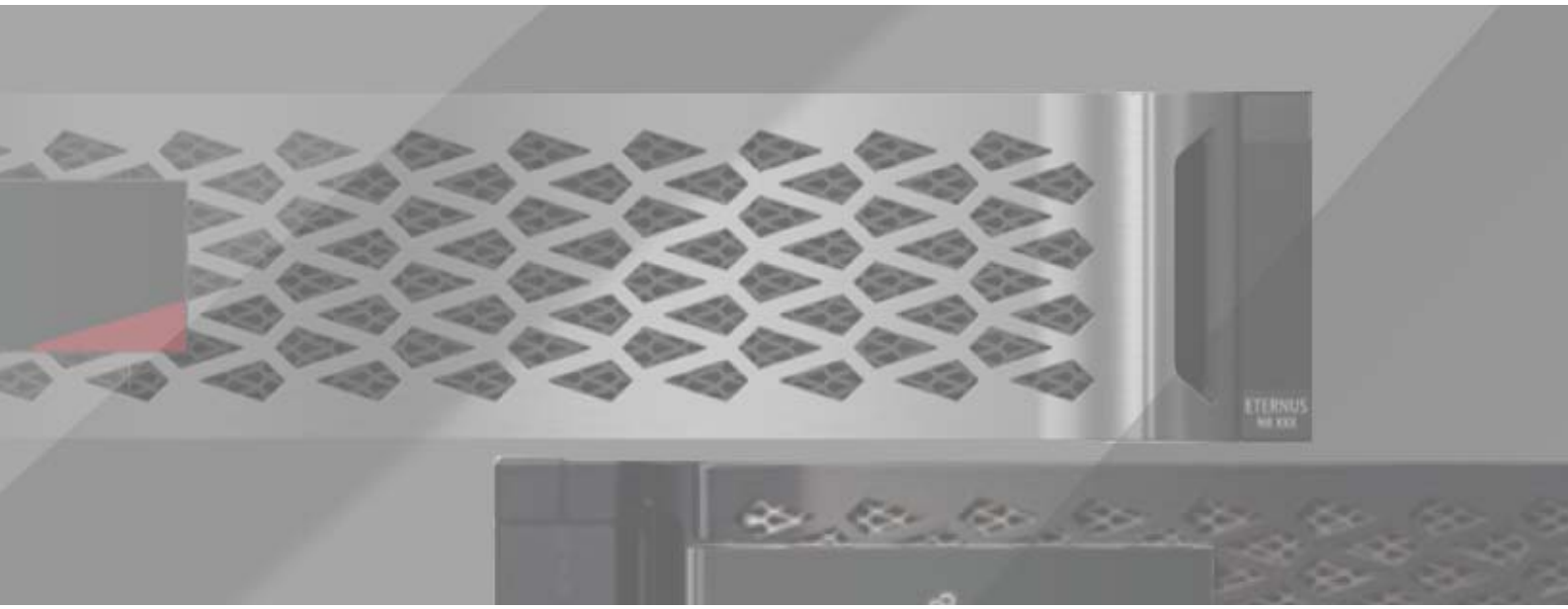
P3AG-6602-01Z0

FUJITSU Storage
ETERNUS AB series オールフラッシュアレイ

# ETERNUS AB series 上の Oracle Database



FUJITSU

# 目次


第 1 章 ETERNUS AB series ストレージシステムの概要 ..... 9
1.1 ETERNUS AB series ハードウェアの概要 ..... 9
1.2 ETERNUS AB3100 にハイブリッド SAS の拡張機能を導入 ..... 10
1.3 SANtricity OS ..... 12
1.3.1 Dynamic Disk Pool ..... 13
1.4 ETERNUS AB series フラッシュアレイのプロビジョニング ..... 15

第 2 章 Oracle Automatic Storage Management ..... 16

第 3 章 高可用性 ..... 18
3.1 ETERNUS AB series システムと SANtricity OS ..... 18
3.2 Oracle HA オプション ..... 19

第 4 章 パフォーマンスの最適化とベンチマーク ..... 20
4.1 ETERNUS AB series の OLTP パフォーマンス ..... 21
4.2 Oracle Automatic Workload Repository とベンチマーク ..... 21
4.3 Oracle AWR およびトラブルシューティング ..... 22
4.4 calibrate_io ..... 23
4.5 SLOB2 ..... 23
4.6 Swingbench ..... 23
4.7 HammerDB ..... 23
4.8 Orion ..... 24

第 5 章 Oracle の一般的な構成 ..... 25
5.1 filesystemio_options ..... 25
5.2 db_file_multiblock_read_count ..... 26
5.3 リドゥブロックサイズ ..... 27
5.4 チェックサムとデータの整合性 ..... 27

第 6 章 サイジング ..... 29
6.1 ETERNUS AB series I/O の概要 ..... 29
6.2 I/O の見積もり ..... 30
6.2.1 新規の OLTP データベースシステム ..... 30
6.2.2 既存の OLTP データベースシステム ..... 31

第 7 章 SANtricity System Manager を使用した ETERNUS AB series のパフォーマンス監視 ..... 32

第 8 章 推奨事項 ..... 34
8.1 ストレージ ..... 34
8.2 Oracle ASM ..... 34
8.3 Linux と NVMe-oF ..... 35

第 9 章 まとめ ..... 36

付録 A Oracle データベースバックアップ用の StorageGRID オブジェクトストレージ ..... 37
A.1 当社のソリューション ..... 37
A.2 クラウドモジュールの構成と StorageGRID へのバックアップを書き込む 5 つのステップ ..... 38
A.3 ベストプラクティス ..... 41
A.4 データベースのバックアップに StorageGRID を使用するメリット ..... 42
A.4.1 複数リージョンクラスタ、単一ネームスペース ..... 42
A.4.2 データの決定性配置 ..... 42
A.4.3 消去コーディング ..... 42

# 図目次


図 1.1 SAS 拡張構成の AB3100 ..... 11
図 1.2 DDP のコンポーネント ..... 13
図 1.3 DDP ドライブの障害 ..... 14
図 2.1 データベースごとの Oracle ASM ディスクグループの配置 ..... 17
図 4.1 OLTP 構成のパフォーマンス結果の比較 ..... 21
図 6.1 Fusion を使用したシステムのサイズ変更と推奨 ..... 31
図 7.1 Logical View を使用したパフォーマンスモニタの例 ..... 32
図 7.2 Physical View を使用したパフォーマンスモニタの例 ..... 33
図 7.3 Applications & Workloads View を使用したパフォーマンスモニタの例 ..... 33

# 表目次


表 1.1 RAID6 における ETERNUS AB series の性能 9
表 1.2 ETERNUS AB series の各モデルの比較 9
表 1.3 AB3100 用ドライブシェルフオプション 11
表 1.4 AB3100 拡張仕様 11
表 6.1 RAID レベルごとの有効容量の比較 30
表 8.1 ストレージチューニングパラメータ 34
表 8.2 Oracle ASM およびインスタンスの設定 34
表 8.3 Linux チューニングパラメータ 35

# はじめに

本書には、ストレージ管理者およびデータベース管理者が ETERNUS AB series ストレージに Oracle を正常にデプロイするために役立つ情報を掲載しています。

Copyright 2022 FUJITSU LIMITED

初版
2022 年 5 月

## 登録商標

本製品に関連する他社商標については、以下のサイトを参照してください。
https://www.fujitsu.com/jp/products/computing/storage/trademark/

本書では、本文中の ™、® などの記号は省略しています。

## 本書の読み方

### 概要

ほとんどの OLTP システムでは、サーバー内のプロセッサ、メモリ、および I/O サブシステムのバランスがよく、パフォーマンスのボトルネックとはみなされません。OLTP 環境におけるパフォーマンスの問題の主な原因は、通常、ストレージの I/O アクティビティに関連しています。既存の HDD ベースストレージシステムの速度は、サーバの処理能力と一致しません。

その結果、強力なプロセッサは頻繁にアイドル状態になり、ストレージ I/O 要求が完了するのを待ちます。この状況は、ユーザーとビジネスの生産性に悪影響を与えます。生産性への影響によって ROI( 投資回収率 ) が低下し、TCO( 総保有コスト ) 全体が上昇します。したがって、ストレージ IOPS のパフォーマンスとレイテンシは、ビジネスにおける戦略的な考慮事項となります。レスポンスタイムの目標が達成され、他のシステムリソース ( プロセッサおよびメモリ ) のパフォーマンスの最適化が実現されていることを確認することが重要です。

ETERNUS AB series フラッシュアレイは、マイクロ秒レベルのレイテンシを要件とするパフォーマンス重視のアプリケーション向けに設計されています。ETERNUS AB series の各フラッシュアレイは、極めて高いパフォーマンスとマイクロ秒レベルのレスポンスタイムを実現し、ビジネスクリティカルなアプリケーションによる迅速な結果の提供とエンドユーザーの操作性の向上を可能にします。このように高い IOPS と超低レイテンシを兼ね備えた ETERNUS AB series フラッシュアレイは、卓越したパフォーマンスを必要とするデータベース駆動型アプリケーションに最適です。

ETERNUS AB series フラッシュアレイは、高パフォーマンスと低レイテンシを提供することで市場をリードしています。AB6100 はコストパフォーマンスに優れたモデルで、AB5100 がこれに続きます。

NVMe は PCIe SSD の業界標準インターフェースとなりました。合理化されたプロトコルコマンドセットと、I/O クロックサイクルの削減により、NVMe は、最大 64K キューとキューあたり最大 64K コマンドをサポートします。これらの属性により、SAS や SATA などの SCSI ベースのプロトコルよりも効率的になります。

NVMe over Fabrics(NVMe-oF) の導入は、インターフェースの特性である低レイテンシと小オーバヘッドに影響を与えることなく、NVMe のスケーラビリティを向上します。ETERNUS AB3100、AB5100、および AB6100 のシステムは、NVMe over RoCE(NVMe/RoCE)、NVMe over InfiniBand(NVMe/IB)、および NVMe over Fibre Channel(NVMe/FC) をサポートしています。

AB3100 と AB6100 の両方で、iSCSI および SCSI ベースのファイバチャネルプロトコル (FCP) と NVMe/FC をサポートします。NVMe/FC を実装するための規格変更はほとんど必要ないため、既存のストレージとともに NVMe/FC を簡単かつシームレスに導入することができ、既存ストレージにも影響をあたえません。また、NVMe/FC は他の FC トラフィックと同じインフラストラクチャコンポーネントを同時に使用できるため、組織に適したペースでワークロードを容易に移行できます。また、NVMe/FC を使用すると、NVMe コマンドと構造を変換せずにエンドツーエンドで効率的に転送できます。

NVMe で構築された AB3100 ストレージアレイは、追加機能として SAS 拡張機能を提供します。コントローラドライブトレイには、最大 240 台の大容量 NL-SAS ドライブまたは 96 台の SAS SSD ドライブと、24 台の NVMe ドライブを搭載できます。Oracle データベースの場合は、NVMe SSD ドライブで高パフォーマンスのデータベースを実行し、データを NL-SAS ドライブにバックアップできます。Oracle データベース管理者 (DBA) が希望する場合は、NL-SAS ドライブまたは SAS SSD ドライブを、Oracle パーティショニングを使用してデータベースのコールド層またはウォーム層としてデプロイできます。

## 対象読者

本書は、お客様の環境への Oracle データベースソリューションを導入する富士通のお客様、パートナー、従業員、およびフィールドスタッフを対象としています。ソリューションのさまざまなコンポーネントに精通していることを前提としています。

## 注意事項

本書では、ETERNUS AB series のオールフラッシュアレイを基盤となるストレージシステムとして使用する、Oracle Automatic Storage Management(Oracle ASM) データベースのストレージレイアウトについて説明します。データベースを ETERNUS AB series ストレージシステムに再配置するか、ETERNUS AB series ストレージシステムで作成して高いパフォーマンスを実現することを前提としています。また、OLTP アプリケーションのパフォーマンスを向上させることを前提としています。

## 関連マニュアル

ETERNUS AB に関連する最新の情報は、以下のサイトで公開されています。
https://www.fujitsu.com/jp/products/computing/storage/manual/

## 本書の表記について

### ■ 本文中の記号

本文中では、以下の記号を使用しています。

| | |
|---|---|
| 注 意 | お使いになるときに注意していただきたいことを記述しています。必ずお読みください。 |
| 備 考 | 本文を補足する内容や、参考情報を記述しています。 |

# 第 1 章
# ETERNUS AB series ストレージシステムの概要

## 1.1 ETERNUS AB series ハードウェアの概要

高い IOPS とミリ秒以下のレスポンスタイムを実現する ETERNUS AB series アレイは、ビジネスクリティカルなアプリケーションの迅速な成果とカスタマーエクスペリエンスの向上を可能にします。

ETERNUS AB series フラッシュアレイは、高い IOPS と極めて低いレイテンシを兼ね備えており、高性能な専用ソリューションを必要とするデータベース駆動型アプリケーションに最適です。

表 1.1 では、RAID6 構成における ETERNUS AB series フラッシュアレイの性能の概要について説明します。

表 1.1　RAID6 における ETERNUS AB series の性能

| 性能比較 | AB2100 | AB3100 | AB5100 | AB6100 |
|---|---|---|---|---|
| SSD 数 | 24 | 24 | 48 | 24 |
| Read IOPS ( 最大 ) | 300K<210µs | 670K<250µs | 1M<250µs | 2M<250µs |
| Write IOPS ( 最大 ) | 45K<160µs | 100K<190µs | 185K<250µs | 340K<190µs |
| Read 帯域幅 | 10GBps | 20GBps | 21GBps | 44GBps |
| Write 帯域幅<br>( キャッシュミラーリングが有効 ) | 3.7GBps | 7GBps | 9GBps | 12.5GBps |
| 低レイテンシ (Read) | 55K<140µs | 150K<200µs | 100K<120µs | 140K<110µs |
| 低レイテンシ (Write) | 45K<160µs | 50K<100µs | 150K<100µs | 200K<80µs |

表 1.2 は、ETERNUS AB series の各モデルの比較結果を示します。

表 1.2　ETERNUS AB series の各モデルの比較

| 製品比較 | AB2100 | AB3100 (*1) | AB5100 | AB6100 |
|---|---|---|---|---|
| 2U/24 拡張シェルフ | 3 | なし | 4 | なし |
| 総ドライブ数 | 96 | 24 | 120 | 24 |
| Raw 容量 | 1.45PB | 367TB | 1.8PB | 367TB |
| コントローラキャッシュオプション | 8GB または 32GB | 16GB | 16GB または 64GB | 32GB または 128GB |

*1:　NVMe コントローラシェルフのみを示しています。SAS 拡張の詳細については、表 1.4 を参照してください。

性能とともに価値を最大化する鍵は、効率を最大化することです。これまで、企業は効率性を犠牲にして、ストレージのオーバープロビジョニングによって、極めて高いパフォーマンスレベルを達成してきました。しかし、状況は変わりつつあります。ETERNUS AB series オールフラッシュアレイは、オーバープロビジョニングを排除することでパフォーマンスと効率性のバランスを取るため、コストを大幅に削減できます。

従来型ドライブ1,000台以上のパフォーマンスを備えたETERNUS AB seriesフラッシュアレイは、ラックスペース、電源、冷却装置を 95% 削減することで、厳しい要件にも対応できます。この機能は、部分的に搭載されたディスクを導入してアプリケーションのパフォーマンスを向上させることに慣れているお客様にとって大きなメリットとなります。

コスト効率に加えて、ETERNUS AB series フラッシュアレイは、アプリケーション効率も向上します。より多くのアプリケーション操作を完了することで、お客様はより効率的になり、より良い結果を得ることができます。

ETERNUS AB series フラッシュアレイには、自動化されたフェイルオーバーを提供する完全冗長の I/O パスがあります。驚くべきことに自動化されたフェイルオーバーは今日利用可能なフラッシュ製品の多くでは提供されていないため、この技術を実装したい企業にとっては絶対的な要件となります。

すべての管理タスクは、ETERNUS AB series アレイがオンライン状態で完全な Read/Write データアクセスを維持している間に実行されます。このアプローチにより、ストレージ管理者は、アプリケーションの I/O を中断することなく、構成の変更やメンテナンスを実行できます。

ETERNUS AB series フラッシュアレイは、データの損失やダウンタイムイベントから保護するために、エンタープライズストレージに共通の高度なデータ保護も提供します。この保護は、Snapshot テクノロジーを使用してローカルで利用でき、また、同期レプリケーションと非同期レプリケーションを使用してリモートからも利用できます（ただし、AB3100、AB6100 を除く）。

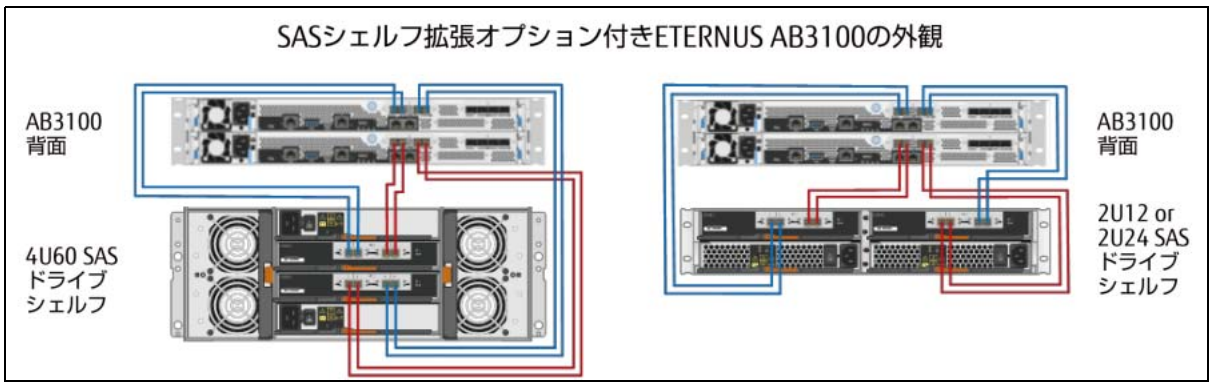
# 1.2　ETERNUS AB3100 にハイブリッド SAS の拡張機能を導入

ETERNUS AB series ストレージアレイは、SAS 拡張シェルフを使用した装置への NL-SAS HDD ドライブの追加を初めてサポートしました。NL-SAS HDD ドライブと SAS SSD ドライブの両方を導入できます。これにより、Oracle データベースのバックアップ場所や、Oracle の階層化機能を使用したコールドデータの HDD ドライブへの移動など、追加の使用例が提供されます。

Oracle データベースでのストレージ階層化技術の使用の詳細は、パーティション化の概要 (oracle.com) を参照してください。

ETERNUS AB3100 は、各コントローラのスロット 1 にオプションの 4 ポート SAS ホストインターフェースカード (HIC) を使用して、拡張機能を提供します。この新しいアーキテクチャは、NVMe シェルフと SAS シェルフを統合したものです。詳細を図 1.1 に示します。

図 1.1　SAS 拡張構成の AB3100




AB3100 コントローラシェルフは NE224 シェルフで 24 台の NVMe SSD ドライブをサポートしますが、コントローラシェルフに拡張ドライブシェルフを追加することで、システムの容量をさらに拡張できます。AB3100 コントローラは、12 Gb SAS3 ドライブシェルフ (DE212C、DE224C および DE460C) とペアにすることができます。これらのシェルフについては、このマニュアルでは詳しく説明しません。表 1.3 に、ドライブシェルフのオプションを示します。

表 1.3　AB3100 用ドライブシェルフオプション

| プロパティ | NE224 | DE212C | DE224C | DE460C |
|---|---|---|---|---|
| フォームファクタ | 2RU | 2RU | 2RU | 4RU |
| ドライブサイズ | 2.5" | 3.5"<br>2.5" ( 取付金具込み ) | 2.5" | 3.5"<br>2.5" ( 取付金具込み ) |
| ドライブタイプ | NVMe SSD | NL-SAS<br>SAS SSD | SAS SSD | NL-SAS<br>SAS SSD |
| 総ドライブ数 | 24 | 12 | 24 | 60 |
| ドライブインターフェース | NVMe | 12Gb SAS | 12Gb SAS | 12Gb SAS |
| 最大シェルフ数 | 1 | 8 | 4 | 4 |

SAS 拡張シェルフを組み合わせると、合計 240 台の NL-SAS ドライブまたは 96 台の SAS SSD を実現できます。表 1.4 に、拡張仕様を示します。

表 1.4　AB3100 拡張仕様

| 仕様 | AB3100 |
|---|---|
| 拡張シェルフを使用しない場合の<br>最大 raw システム容量 (SSD 24 台を想定 ) | 367TB (15.3TB　SSD x 24 台 ) |
| 拡張シェルフを使用した場合の<br>最大 raw システム容量 (24 台の NVMe SSD と 240 台の NL-SAS ドライブを想定 ) | 4,687TB<br>4,320TB (18TB NL-SAS x 240 台 ) + 367 TB (15.3TB SSD x 24 台 ) |

| 仕様 | AB3100 |
|---|---|
| オプションのドライブシェルフ拡張 | • スロット 1 だけに 12Gb SAS を搭載 ( コントローラあたり 4 ポート )<br>• サポートされる NL-SAS ドライブの最大拡張数 : DE212C および DE460C シェルフを混在させる場合は、DE212C シェルフだけが使用され、8 つの DE212C シェルフが使用できる場合を除き、NL-SAS ドライブスロットは合計 240 個まで、拡張シェルフは 4 個までとしてください。たとえば、4 つの DE460C シェルフ、8 つの DE212C シェルフ、または 2 つの DE460C シェルフ +2 つの DE212 シェルフです。<br>• サポートされる SAS SSD ドライブの最大拡張数 : DE212C、DE224C、および DE460C シェルフを混在させる場合は、DE212C シェルフだけを使用し、DE212C シェルフを 8 つ使用できる場合を除き、SAS SSD ドライブスロットは合計で 96 個まで、拡張シェルフは 4 個までとしてください。たとえば、1 つの DE460C シェルフ +1 つの DE224C シェルフ +1 つの DE212C シェルフ、または 4 つの DE224C シェルフ、または 8 つの DE212C シェルフです。<br>**注 意**<br>• 10k SAS ドライブはサポートされていません。<br>• NVMe ドライブを含む 2 番目のエンクロージャへの拡張はサポートされていません。 |

# 1.3　SANtricity OS

ETERNUS AB series は、ブラウザベースのアプリケーションである SANtricity System Manager で管理されます。SANtricity System Manager はコントローラーに組み込まれています。

アレイ上にボリュームグループを作成する最初のステップとして、構成時に保護レベルを割り当てます。割り当てた保護レベルは、ボリュームグループを形成するために選択したディスクに適用されます。ETERNUS AB series は、Dynamic Disk Pool(DDP) と RAID レベル 0、1、5、6、および 10 をサポートします。DDP は、このマニュアルで説明するすべての設定に使用されています。

ストレージのプロビジョニングを簡素化するために、SANtricity System Manager には自動構成機能が用意されています。構成ウィザードは、アレイ上の使用可能なディスク容量を分析します。次に、容量要件、ホットスペア、およびウィザードで指定されたその他の条件を満たしながら、アレイのパフォーマンスとフォールトトレランスを最大化するディスクを選択します。

SANtricity Unified Manager および SANtricity System Manager の詳細については、富士通マニュアルサイトを参照してください。

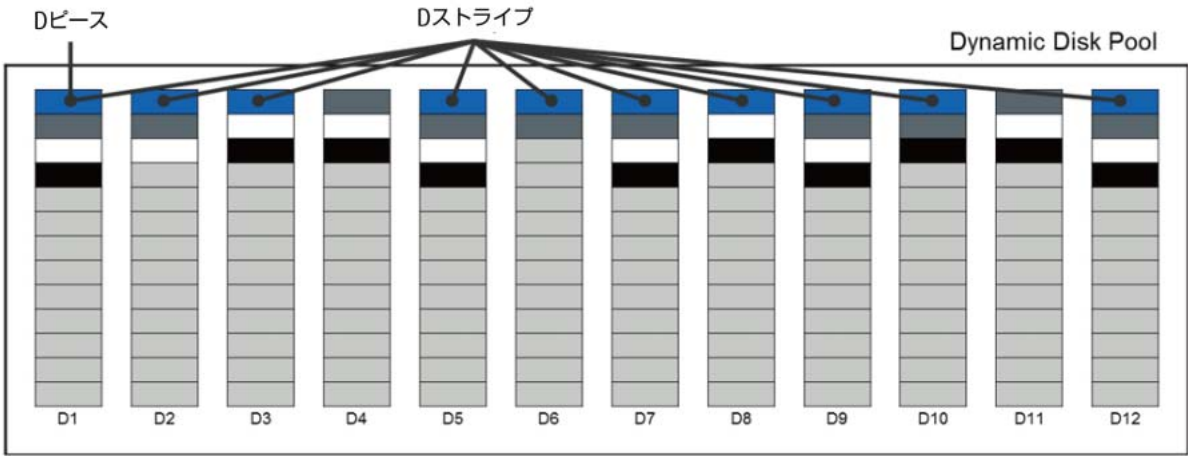
## 1.3.1　Dynamic Disk Pool

7 件の特許が出願中の DDP 機能は、データ、スペア容量、保護情報をドライブのプール全体に動的に分散します。これらのプールの範囲は、最小で 11 台のドライブから、最大でシステムがサポートするすべてのドライブまでです。ストレージ管理者は、単一のプールを作成するだけでなく、従来のボリュームグループと DDP または複数のプールを混在させることができ、これまでにないレベルの柔軟性を実現できます。

プールは、複数の下位レベルの要素で構成されます。最初の要素は D ピースです。D ピースは 512MB の一続きのセクションで、物理ドライブ内の 4,096 個の 128KB セグメントで構成されます。システムは、プール内の選択されたドライブから、インテリジェントな最適化アルゴリズムを使用して 10 個の D ピースを選択します。10 個の関連する D ピースを合わせると 1 つの D ストライプと見なされ、有効容量は 4GB になります。D ストライプの内容は RAID6(8+2) のシナリオと似ています。この場合、8 つの基本セグメントにユーザーデータ、1 つのセグメントにユーザーデータセグメントから計算されたパリティ (P) 情報、1 つのセグメントに RAID6 で定義された Q 値が含まれます。

その後、プール内の最大許容ボリュームサイズまで定義済みボリュームサイズを満たすために、必要に応じて複数の 4GB D ストライプアグリゲートからボリュームが作成されます。図 1.2 は、これらのデータ構造間の関係を示しています。

図 1.2　DDP のコンポーネント

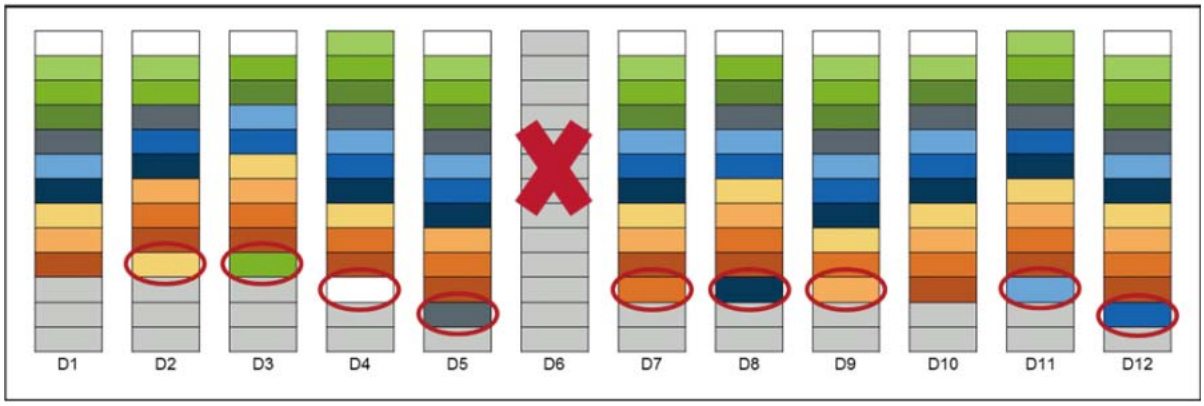



プールのもう 1 つの大きなメリットは、孤立状態の専用ホットスペアを使用するのではなく、プールに統合された保存容量があり、ドライブ障害に備えた再構築領域を提供できることです。このアプローチにより、個々のホットスペアを計画または管理する必要がなくなり、管理が合理化されます。また、従来のホットスペアとは対照的に、必要に応じて再構築の時間を大幅に改善し、再構築中のボリュームパフォーマンスを向上させます。

プール内のドライブに障害が発生すると、RAID6 で通常使用されるのと同じメカニズムを使用して、障害が発生したドライブの D ピースがプール内に存在する他のすべてのドライブに再構築されます。このプロセスでは、コントローラフレームワーク内部のアルゴリズムによって、1 台のドライブに同じ D ストライプからの D ピースが 2 つ含まれていないことが検証されます。個々の D ピースは、選択したドライブ上で最も低い論理ブロックアクセス (LBA) 範囲で再構築されます。

図 1.3 は、ドライブ 6(D 6) に障害が発生した場合の例を示します。次に、そのディスク上に存在していた D ピースが、プール内の他の複数のドライブに同時に再作成されます。この処理には複数のディスクが関与しているため、この状況による全体的なパフォーマンスへの影響が軽減され、処理の完了に必要な時間が大幅に短縮されます。

図 1.3　DDP ドライブの障害




プール内で複数のディスク障害が発生した場合、データ可用性のリスクを最小限に抑えるために、D ピースが 2 つ欠落している D ストライプの再構築が優先されます。これらの重大な影響を受けた D ストライプが再構成された後、必要なデータの残りが再構成されます。

コントローラのリソース割り当ての観点から、DDP にはユーザーが変更可能な再構築優先順位 2 つ用意されています。

- Degraded reconstruction priority( デグレードの再構築優先度 ) は、影響を受けた D ストライプに再構築が必要な D ピースが 1 つしかない場合に適用される優先度です ( 値のデフォルトは High です )。
- Critical reconstruction priority( クリティカルな再構築優先度 ) は、D ストライプに再構築が必要な D ピースが 2 つある場合に適用される優先度です ( 値のデフォルトは Highest です )。

大規模なドライブプールで同時に 2 つのディスクに障害が発生した場合、2 つの D ピースの再構築が必要な状況に陥る D ストライプの数はそれほど多くはありません。前述したように、これらの重要な D ピースは、最初に最も高い優先順位で識別および再構築されます。そうすることで、プールはすぐにデグレード状態に戻り、それ以上のドライブ障害に耐えることができるようになります。

DDP は、再構築時間を短縮し、優れたデータ保護を提供するだけでなく、障害が発生した状態でのベースボリュームのパフォーマンスを、従来のボリュームグループのパフォーマンスと比べて大幅に向上させることができます。

DDP の詳細については、富士通マニュアルサイトの「ETERNUS AB/HB series SANtricity OS Dynamic Disk Pool 機能の説明とベストプラクティス」を参照してください。

# 1.4　ETERNUS AB series フラッシュアレイのプロビジョニング

SANtricity DDP テクノロジーは、ストレージ管理者に RAID 管理の合理化、データ保護の向上、あらゆる条件下での予測可能なパフォーマンスの維持を提供します。DDP は、ETERNUS AB series のドライブプール全体にデータ、保護情報、スペア容量を均等に分散し、セットアップを効率化して使用率を最大化します。この次世代テクノロジーは、ドライブ障害によるパフォーマンスへの影響を最小限に抑え、従来の RAID よりも最大 8 倍高速にシステムを最適な状態に戻すことができます。再構築時間が短縮され、再構築を優先する特許取得済みのテクノロジーにより、DDP は複数のディスク障害のリスクを大幅に軽減し、従来の RAID では実現できないレベルのデータ保護を提供します。

SANtricity ソフトウェアでは、ストレージをオンラインに保ったまますべての管理タスクを実行できるため、データへの常時アクセス ( 読み取り、書き込み ) が維持されます。ストレージ管理者は、接続されたホストへの I/O を中断することなく、構成の変更、メンテナンス、ストレージ容量の拡張を行うことができます。SANtricity ソフトウェアのオンライン機能には、次のものがあります。

- 動的なボリューム拡張により、管理者は既存のボリュームの容量を拡張できます。
- 動的なセグメントサイズの移行により、管理者は特定のボリュームのセグメントサイズを変更できます。
- 動的な RAID レベルの移行では、データを再配置しなくても、既存のドライブ上の RAID グループの RAID レベルを変更できます。サポートされる RAID レベルは、0、1、5、6、および 10 です。
- 無停止でのコントローラファームウェアのアップグレードは、データアクセスを中断することなくサポートされます。

ETERNUS AB series フラッシュアレイのプロビジョニングの詳細については、富士通マニュアルサイトを参照してください。ボリュームの作成とデプロイが繰り返される場合は、REST API コマンドを使用してこのタスクを自動化できます。

## ■ ベストプラクティス

富士通では、すべてのストレージ構成に Dynamic Disk Pool を使用することを推奨します。

# 第 2 章
# Oracle Automatic Storage Management

Oracle Automatic Storage Management(Oracle ASM) では、クラスタファイルシステムとボリューム管理の統合機能が提供されるため、サードパーティのボリューム管理ツールが不要になり、アーキテクチャ全体の複雑さが軽減されます。

Oracle ASM の主な機能は次のとおりです。

- ファイルとボリュームの自動管理
- RAW I/O のパフォーマンスを備えたデータベースファイルシステム
- データの自動分散とストライピング
- 外部 ( アレイベース ) のデータ保護と 2 方向および 3 方向のミラー保護を選択可能
- ミラーデータのどのコピーを優先的に使用するかを制御

これらの機能により、Oracle ASM は、サードパーティのファイルシステムおよびボリューム管理ソリューションに代わって、データベースの作成または配置、ディスク領域の使用の管理などのデータベース記憶域管理タスクを行います。Oracle サーバのホストでのボリューム管理タスクは、使い慣れた create、alter、drop などの SQL 文を使用して実行できます。これにより、データベースのストレージプロビジョニングに関するデータベース管理者の作業が簡略化されます。ロードバランシングは、I/O ワークロードが使用可能なすべてのドライブリソースを使用できるようにすることで、パフォーマンスのボトルネックを回避します。

ETERNUS AB series ストレージシステムは、基礎となるボリュームグループのすべての Solid State Drive (SSD) 間で I/O のロードバランシングを自動的に行います。1 つのボリュームグループ内に配置されたすべての LUN は、ボリュームグループのすべての SSD をバランスよく使用できます。Oracle ASM では、64MB のストライプサイズに基づいて各データファイルの内容をディスクグループの記憶域プール全体に均等に分散することで、Oracle ASM ディスクグループ内のすべての LUN またはファイル間で I/O のロードバランシングをさらに強化します。これにより、ホスト層とネットワーク層で使用可能な SCSI デバイスを介して、パフォーマンスが均一になります。

ロードバランシングを Oracle ASM とともに使用すると、複数の LUN およびファイルシステムデータで共通のドライブを共有できます。この機能により、各 Oracle ASM ディスクグループに必要な LUN の数が削減され、パフォーマンスを低下させることなく管理性が向上します。これにより、トランザクションの多いデータベース環境で最適な Read および Write パフォーマンスが得られます。

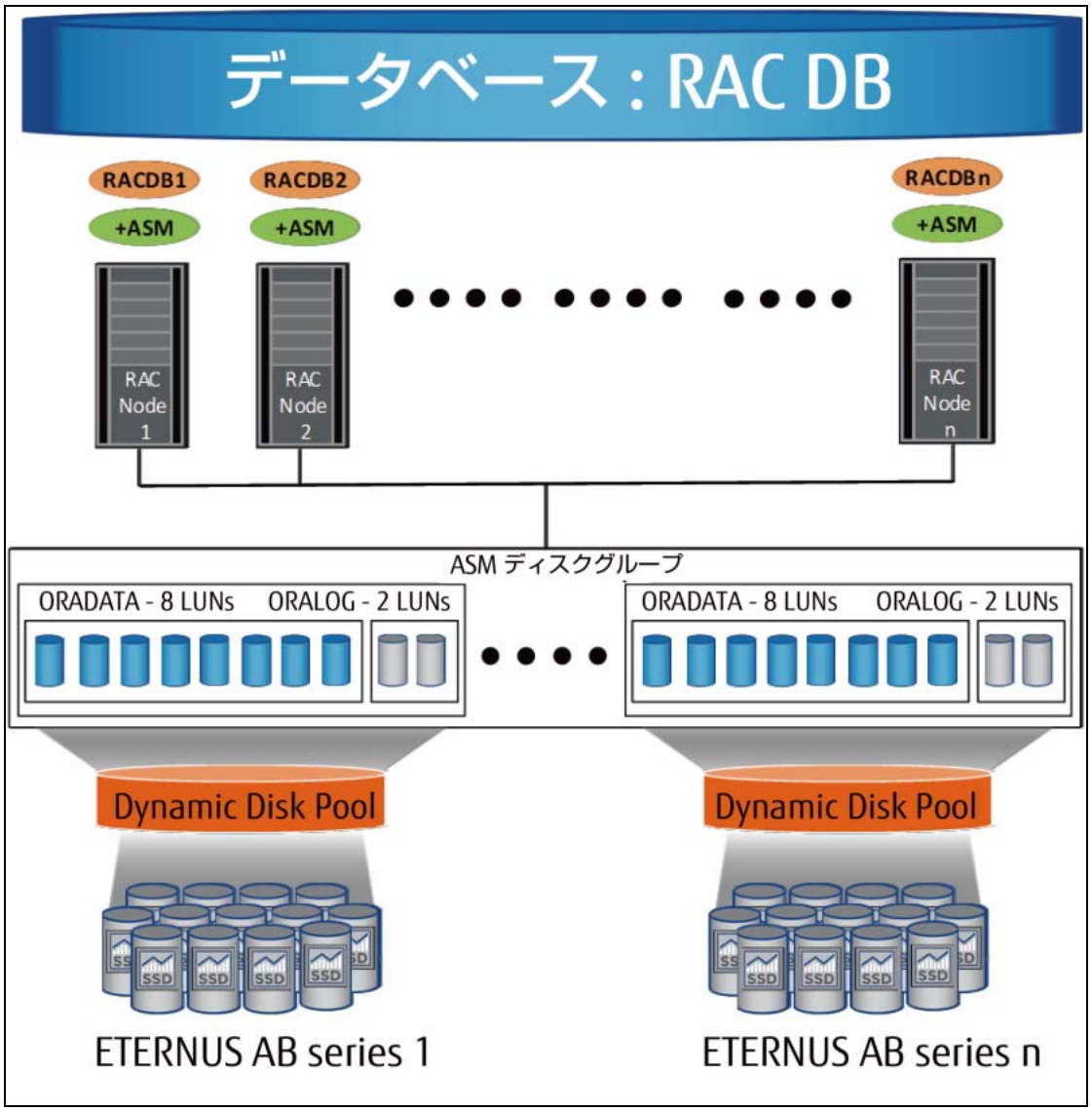
## ■ ベストプラクティス

Oracle ASM ディスクグループを ETERNUS AB series ストレージに作成する場合は、データディスクグループに対して DDP ごとに 8 個以上の LUN を構成することを推奨します。複数の LUN を構成すると、ディスクグループに対する I/O の同時実行性が最大化され、ホストドライバスタック (OS、ホストバスアダプタ (HBA) ドライバ、またはマルチパスドライバ ) の LUN ごとの制限を克服できます。このプロビジョニングは、最初は不要であっても、同じボリューム内の LUN が後でディスクグループに追加された場合に発生する不要なデータ移動を回避できます。

図 2.1 に、複数の ETERNUS AB series オールフラッシュアレイを使用して、マルチノード RAC データベース用の Oracle RAC データベースを構成する方法を示します。各 ETERNUS AB series で使用可能なすべてのドライブを使用して単一の DDP が作成され、そこからデータボリュームとログボリュームが作成されます。ORADATA および ORALOG LUN は、SANtricity System Manager を使用して ETERNUS AB series アレイ上でプロビジョニングされ、Oracle RAC ノードに提示されます。データベースおよびホストの観点から単一のファイルシステムを提供するために、すべての LUN にまたがる単一の ASM ディスクグループを作成できます。

この構成によって、すべてのアレイ上のすべてのストレージ LUN とボリュームにワークロードが均等に分散されます。ETERNUS AB series ではすでにアレイレベルでの保護が提供されているため、ASM ディスクグループには外部冗長性が採用されました。ストレージアレイに十分な容量がプロビジョニングされている場合は、ASM ミラーリングを選択してデータ保護を強化することもできます。

図 2.1　データベースごとの Oracle ASM ディスクグループの配置



# 第 3 章
# 高可用性

## 3.1 ETERNUS AB series システムと SANtricity OS

ETERNUS AB series ストレージシステムは、高い信頼性と高可用性を実現するように設計されており、次のような機能を備えています。

- 自動 I/O パスフェイルオーバーを備えたデュアルアクティブコントローラ
- RAID レベル 0、1、5、6、10、または DDP
- ホットスワップ対応の冗長コントローラ、ディスク、電源装置、およびファン
- グローバルホットスペアを使用した自動ドライブフェイルオーバー検出と再構築
- バッテリのバックアップとメモリへのデステージ機能を備えたミラーデータキャッシュ
- 無停止でのコントローラファームウェアアップグレード
- ドライブの健全性を積極的に監視
- 自動パリティチェックと修正を備えたバックグラウンドメディアスキャン

すべてのコンポーネントは完全な冗長性を備えており、システムの電源を切ったり、動作を停止したりすることなく、コンポーネントを交換できます。冗長コンポーネントには、コントローラ、ディスク、電源装置、およびファンが含まれます。ETERNUS AB series の電源装置は、80-PLUS 認証電源です。ETERNUS AB series フラッシュアレイは、あらゆる状況でデータを保護するために設計されたいくつかの機能を備えています。さまざまな冗長性レベルで複数の RAID レベルを使用できます。接続が失われると、あるパスから別のパスへの自動フェイルオーバーもシステムに含まれます。シェルフ内では、各ドライブが各コントローラに接続されるため、内部接続の問題も迅速に解決できます。ETERNUS AB/HB series で作成したボリュームは、作成後すぐにホスト I/O に使用できます。また、I/O を中断することなく、多くのプロパティを変更できます。

データを保護する ETERNUS AB series フラッシュアレイのその他の機能には、コントローラキャッシュのミラーリングとバックアップがあります。システムの動作中に電源が失われた場合は、オンボードバッテリがキャッシュメモリから内部コントローラフラッシュにデータをデステージして、電源が復旧したときにデータを使用できるようにします。RAID アルゴリズムを使用すると、ドライブに障害が発生した場合に失われたデータをシステムで再作成できます。また、RAID パリティを使用してデータを確認し、読み取り不能セクタにヒットした場合は再構築を続行することもできます。

システムは、データを常に保護する他のタスクをバックグラウンドで実行します。オプションのメディアスキャン機能では、現在どのホストからもアクセスされていないセクタについても、不整合を検索します。ETERNUS AB series アレイは、SSD の消耗を積極的に追跡し、寿命が近づいているドライブにフラグを立てます。すべてのタイプの診断データは、必要に応じて富士通サポートが後で使用できるように定期的に収集されます。

すでに説明したように、ETERNUS AB series アレイには、信頼性と可用性のための様々な機能が用意されています。さらに、SANtricity ソフトウェアを使用すると、可用性を最大限に高めることができます。たとえば、SANtricity は以下のタスクを実行します。

- 高速かつ効率性に優れた Snapshot テクノロジーを活用
- 数秒でデータを保護
- 変更されたブロックのみを格納することでフラッシュの消費量を削減
- 堅牢なディザスタリカバリ保護を提供

- 同期ミラーリングをサポートし、データ消失ゼロのコンテンツ保護を実現
- 非同期ミラーリングよって遠隔地の保護とコンプライアンスをサポート
- 柔軟な保護によって ROI を最大化
- コストとパフォーマンスのニーズに基づいて、フラッシュ、ニアライン SAS(NL-SAS)、または複数のリカバリターゲットの混在をサポート
- 予算を超えることなく迅速に処理

## 3.2　Oracle HA オプション

Oracle には、様々なデータベース構成に対応する多数の高可用性オプションが用意されています。さまざまなオプションの説明は、ここでできます。

# 第 4 章
# パフォーマンスの最適化とベンチマーク

データベースストレージのパフォーマンスの正確な試験を行うのは難しい課題です。IOPS とスループットを理解する必要があります。このテーマでは、フォアグラウンド I/O 操作とバックグラウンド I/O 操作の違い、データベースに対する待機時間の影響、ストレージパフォーマンスにも影響を与える多数の OS およびネットワーク設定についても理解する必要があります。また、ストレージ以外のデータベースタスクも考慮する必要があります。ストレージパフォーマンスは性能の制限要因ではなくなったため、ストレージパフォーマンスを最適化しても有益なメリットは得られません。

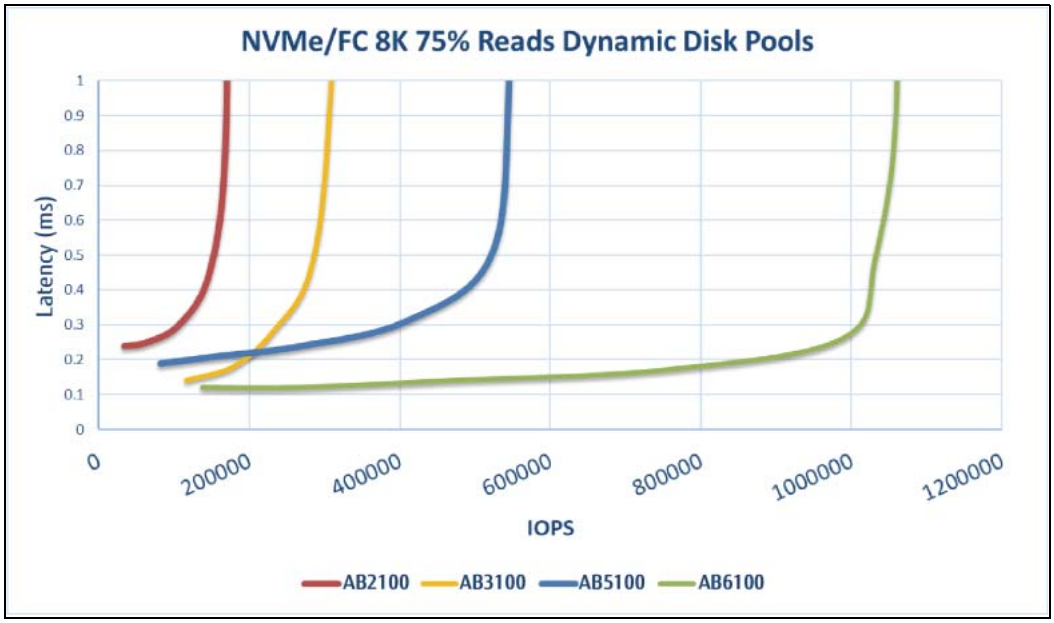
ETERNUS AB series のようなオールフラッシュアレイを選択する場合は、ほかにも考慮事項があります。

- 75:25 の Read:Write 比により、AB6100 は 300µs のレイテンシで約 800K のランダムデータベース IOPS を実現できます。これは、ほとんどのデータベースの現在のパフォーマンス要件をはるかに超えているため、期待性能を予測することは困難です。ストレージの大部分はボトルネックとして除外されます。
- ネットワーク帯域幅は、パフォーマンス制限の原因としてますます一般的になっています。たとえば、回転ディスクソリューションは、I/O 遅延が非常に長いため、データベースパフォーマンスのボトルネックになることがよくあります。オールフラッシュアレイによって遅延の制限が取り除かれると、問題がネットワークに移ることがよくあります。これは、実際のネットワーク接続を視覚化するのが難しい仮想化環境やブレードシステムで特に顕著です。帯域制限によりストレージシステム自体を完全に利用できない場合は、パフォーマンステストが複雑になる可能性があります。
- オールフラッシュアレイのパフォーマンスを、回転ディスクを含むアレイと比較することは、オールフラッシュアレイのレイテンシが大幅に改善されているため、一般に不可能です。通常、試験結果に意味はありません。
- データベースはストレージ I/O によって制限されないため、ピーク時の IOPS パフォーマンスをオールフラッシュアレイと比較することは、あまり有用な試験はありません。たとえば、あるアレイが 500K のランダム IOPS を維持でき、別のアレイが 300K を維持できるとします。データベースがその時間の 99% を CPU 処理に費やしている場合、現実にはこの違いは関係ありません。ワークロードがストレージアレイの機能をフルに活用することはありません。これとは対照的に、統合プラットフォームでは、ストレージアレイがそのピーク時の性能に対応する必要があるため、ピーク時の IOPS 機能が非常に重要になる可能性があります。
- ストレージテストでは、常にレイテンシと IOPS を考慮してください。市場に出回っているストレージアレイの多くは、IOPS が非常に高いと主張していますが、そのようなレベルでは、レイテンシが原因で IOPS が使用できなくなります。オールフラッシュアレイの一般的なターゲットは、1ms マークです。より有益な試験は、最大可能 IOPS を測定することではなく、平均遅延が 1ms を超える前にストレージアレイが維持できる IOPS の数を判断することです。

# 4.1　ETERNUS AB series の OLTP パフォーマンス

図 4.1 は、一般的な OLTP ワークロードにおいて DDP で構成した ETERNUS AB series( エントリーモデルの AB2100 (FC)、ミッドレンジモデルの AB5100、エンド・ツー・エンド NVMe モデルの AB3100 および AB6100) の性能結果を示します。試験は、プールで 24 台の SSD を使用し、AB3100 と AB6100 では FC ホストプロトコルまたは NVMe/FC を使用し、ワークロードを Read75% と Write25% に設定したアレイで実施されました。レイテンシが約 300µs のとき、AB2100 は 100,000IOPS、AB5100 は 400,000IOPS を実現しますが、NVMe ベースの AB3100 と AB6100 はそれぞれ 240,000IOPS と 1,000,000IOPS 以上を実現します。

図 4.1　OLTP 構成のパフォーマンス結果の比較




# 4.2　Oracle Automatic Workload Repository とベンチマーク

Oracle パフォーマンス比較の絶対基準は、Oracle　Automatic Workload Repository(AWR) レポートです。

AWR レポートには複数のタイプがあります。ストレージの観点から見ると、`awrrpt.sql` コマンドを実行して生成されるレポートは、特定のデータベースインスタンスを対象とし、レイテンシに基づいてストレージ I/O イベントを細分化する詳細なヒストグラムが含まれているため、最も包括的で有意義です。

2 つのパフォーマンスアレイを比較するには、各アレイで同じワークロードを実行し、ワークロードを正確にターゲットとする AWR レポートを作成することが理想です。実行時間が非常に長いワークロードの場合は、開始時間と停止時間を含む経過時間を含む単一の AWR レポートを使用できますが、AWR データを複数のレポートに分割することを推奨します。たとえば、バッチジョブが午前 0 時から午前 6 時まで実行された場合、午前 0 時から午前 1 時まで、午前 1 時から午前 2 時までなど、1 時間ごとの AWR レポートを作成します。

それ以外の場合は、非常に短いクエリを最適化する必要があります。最良の選択肢は、クエリの開始時に作成された AWR スナップショットと、クエリの終了時に作成された 2 つ目の AWR スナップショットに基づく AWR レポートです。データベースサーバは、分析中のクエリのアクティビティを不明確にするバックグラウンドアクティビティを最小限に抑えるため、サイレントにする必要があります。

備 考

AWR レポートが使用できない場合は、Oracle Statspack レポートを使用することを推奨します。このレポートには、AWR レポートとほとんど同じ I/O 統計が含まれます。

# 4.3　Oracle AWR およびトラブルシューティング

AWR レポートは、パフォーマンスの問題を分析するための最も重要なツールでもあります。

ベンチマークと同様に、パフォーマンスのトラブルシューティングでは、ワークロードを正確に測定する必要があります。可能な場合は、パフォーマンスの問題を富士通サポートセンターに報告するとき、または新しいソリューションについて富士通またはパートナーアカウントチームと作業するときに、AWR データを提供してください。

AWR データを提供する場合は、次の要件を考慮してください。

- `awrrpt.sql` コマンドを実行してレポートを生成します。テキストまたは HTML のいずれかで出力されます。
- Oracle RAC が使用されている場合は、クラスタ内のインスタンスごとに AWR レポートを生成します。
- 問題が発生した特定の時刻をターゲットにします。AWR レポートの最大許容経過時間は、通常 1 時間です。問題が 1 時間以上続く場合、またはバッチジョブなどの 1 時間以上かかる操作を伴う場合は、分析対象期間全体をカバーする複数の 1 時間 AWR レポートを提供します。
- 可能な場合は、AWR スナップショット間隔を 15 分に調整します。この設定により、より詳細な解析を実行できます。15 分間隔ごとのレポートを提供するためには `awrrpt.sql` コマンドの追加実行も必要です。
- 実行時間が非常に短いクエリの問題の場合は、操作の開始時に作成された AWR スナップショットと、操作の終了時に作成された 2 つ目の AWR スナップショットに基づいて AWR レポートを作成します。データベースサーバは、分析中の操作のアクティビティを不明確にするバックグラウンドアクティビティを最小限に抑えるため、サイレントにする必要があります。
- パフォーマンスの問題が特定の時間に報告され、他の時間に報告されない場合は、比較に適したパフォーマンスを示す追加の AWR データを提供します。

## 4.4　calibrate_io

`calibrate_io` コマンドは、ストレージシステムのテスト、比較、またはベンチマークには決して使用しないでください。Oracle のマニュアルに記載されているように、この手順ではストレージの I/O 機能を調整します。

キャリブレーションはベンチマークとは異なります。このコマンドの目的は、データベース操作を調整する I/O を発行し、ホストに発行される I/O のレベルを最適化することによって、その効率を向上させることです。`calibrate_io` 操作によって実行される I/O のタイプは、実際のデータベースユーザー I/O を表していないため、結果は予測できず、再現性もないことがよくあります。

## 4.5　SLOB2

Silly Little Oracle Benchmark (SLOB2) は、データベースのパフォーマンスを評価するための推奨ツールとなりました。Kevin Closson によって開発された本ツールは、こちらから入手できます。インストールと構成には数分かかり、実際の Oracle データベースを使用してユーザーが定義可能なテーブル領域に I/O パターンを生成します。I/O でオールフラッシュアレイを飽和状態にすることができる、数少ない有効なテストオプションの 1 つです。また、非常に低いレベルの I/O を生成して、IOPS は低くてもレイテンシの影響を受けやすいストレージワークロードをシミュレートする場合にも役立ちます。

## 4.6　Swingbench

Swingbench はデータベースのパフォーマンスをテストするのに役立ちますが、ストレージに負荷をかけるような方法で Swingbench を使用することは非常に困難です。富士通は、どの ETERNUS AB series アレイでも大きな負荷になるような十分な I/O を実現する Swingbench のテストを行っていません。場合によっては、Order Entry Test(OET) を使用してレイテンシの観点からストレージを評価できます。これは、データベースに特定のクエリに対する既知のレイテンシ依存がある場合に役立ちます。ホストとネットワークがオールフラッシュアレイの潜在レイテンシを実現するように正しく設定されていることを確認する必要があります。

## 4.7　HammerDB

HammerDB は、TPC-C や TPC-H などのベンチマークをシミュレートするデータベーステストツールです。テストを適切に実行するために十分な大きさのデータセットを構築するには多くの時間がかかりますが、OLTP およびデータウェアハウスアプリケーションのパフォーマンスを評価するための効果的なツールになります。

# 4.8 Orion

Oracle Orion ツールは Oracle 9 で一般的に使用されていましたが、さまざまなホストオペレーティングシステムの変更との互換性を確保するためにメンテナンスされていませんでした。OS およびストレージ構成との互換性がないため、Oracle 10 または Oracle 11 ではほとんど使用されません。

Oracle はこのツールを書き直し、デフォルトで Oracle 12c とともにインストールされるようにしました。この製品は改善されており、実際の Oracle データベースで使用されるのと同じ call の多くを使用しますが、Oracle で使用されるのと厳密に同じコードパスまたは I/O 動作は使用しません。たとえば、ほとんどの Oracle I/O は同期的に実行されます。つまり、I/O 操作はフォアグラウンドで完了するため、I/O が完了するまでデータベースは停止します。ストレージシステムにランダム I/O をフラッディングするだけでは、実際の Oracle I/O を再現することはできず、ストレージアレイを比較したり、構成変更の影響を測定したりする直接的な方法を提供することはできません。

ただし、特定のホスト / ネットワークストレージ構成の最大可能パフォーマンスの一般的な測定や、ストレージシステムの稼働状態の測定など、Orion にはいくつかの使用例があります。慎重なテストを行って、IOPS、スループット、およびレイテンシを考慮したパラメータを使用した現実的なワークロードを再現することができれば、ストレージアレイの比較や構成変更の影響を評価することができる Orion テストを考案できます。

# 第 5 章
# Oracle の一般的な構成

以下のパラメータは、一般にすべての構成に適用できます。

## 5.1 filesystemio_options

Oracle 初期化パラメータ `filesystemio_options` は、非同期および直接 I/O の使用を制御します。通説とは異なり、非同期 I/O と直接 I/O は相互に排他的ではありません。富士通では、お客様の環境でこのパラメータの設定ミスが頻繁に発生することを確認しています。この設定ミスが多くのパフォーマンス問題の直接の原因となっています。

非同期 I/O とは、Oracle I/O 操作をパラレル化できることを意味します。様々な OS で非同期 I/O が使用可能になる前は、ユーザーは多数の dbwriter プロセスを構成し、サーバプロセスの構成を変更していました。非同期 I/O では、OS 自体がデータベースソフトウェアにかわって I/O を効率的かつパラレルに実行します。このプロセスによってデータが危険にさらされることはなく、Oracle redo ロギングなどの重要な操作は同期的に実行されます。

直接 I/O では、OS バッファキャッシュがバイパスされます。UNIX システム上の I/O は通常、OS バッファキャッシュを介して処理されます。これは、内部キャッシュを保持しないアプリケーションには有用ですが、Oracle は SGA 内に独自のバッファキャッシュを持っています。ほとんどの場合、OS バッファキャッシュに依存するよりも、直接 I/O を有効にしてサーバーRAM を SGA に割り当てる方が適切です。Oracle SGA はメモリをより効率的に使用します。また、OS バッファを経由して I/O が流れると、追加処理が発生してレイテンシが長くなります。低レイテンシが重要な要件の場合に大量の Write I/O が発生すると、特に顕著にレイテンシが増加します。

`filesystemio_options` のオプションは次のとおりです。

- **async**
  Oracle は、処理のために OS に I/O 要求を発行します。このプロセスにより、Oracle は I/O の完了を待たずに他の作業を実行できるため、I/O の並列化が向上します。
- **directio**
  Oracle は、ホスト OS キャッシュを介して I/O をルーティングするのではなく、物理ファイルに対して I/O を直接実行します。
- **none**
  Oracle では、同期 I/O とバッファ I/O が使用されます。この構成では、共有サーバプロセスと専用サーバプロセスの選択、および dbwriter の数がより重要になります。
- **setall**
  Oracle では、非同期 I/O と直接 I/O の両方が使用されます。

ほとんどの場合、`setall` の使用が最適ですが、次の点を考慮してください。

- 一部のお客様では、特に以前の Red Hat Enterprise Linux 4 (RHEL4) リリースで、非同期 I/O の問題が発生していました。しかしこれらの問題は報告されなくなり、非同期 I/O は現在のすべての OS で安定しています。

- データベースでバッファ I/O が使用されている場合は、直接 I/O への切り替えによって SGA サイズの変更が保証されることもあります。バッファ I/O を無効にすると、ホスト OS キャッシュがデータベースにもたらすパフォーマンス上のメリットがなくなります。SGA に RAM を再追加すると、この問題が修復されます。その結果、I/O パフォーマンスが向上します。
- ほとんどの場合、OS バッファキャッシュよりも Oracle SGA に RAM を使用する方が適していますが、最適な値を判断できない場合があります。たとえば、断続的にアクティブな Oracle インスタンスが多数存在するデータベースサーバでは、SGA サイズが非常に小さいバッファ I/O を使用することを推奨します。この構成により、実行中のすべてのデータベースインスタンスで OS 上の残りの空き RAM を柔軟に使用できます。これは非常にまれな状況ですが、一部のお客様環境で確認されています。

備 考

`filesystemio_options` パラメータは、DNFS および ASM 環境では無効です。DNFS または ASM を使用すると、非同期 I/O と直接 I/O の両方が自動的に使用されます。

### ■ ベストプラクティス

`filesystemio_options` を `setall` に設定することを推奨しますが、状況によっては、ホストのバッファキャッシュ失われることによって Oracle SGA の増加が必要になる場合があることに注意してください。

## 5.2　db_file_multiblock_read_count

`db_file_multiblock_read_count` パラメータは、シーケンシャル I/O 中に Oracle が 1 回の操作で読み取る Oracle データベースブロックの最大数を制御します。ただし、このパラメータは、一部またはすべての読み取り操作中に Oracle が読み取るブロック数に影響するわけではなく、ランダム I/O にも影響しません。影響を受けるのはシーケンシャル I/O のみです。

このパラメータは設定しないことを推奨します。設定しないことにより、データベースソフトウェアは最適な値を自動的に設定できます。つまり、このパラメータは I/O サイズが 1MB になるように設定するということです。たとえば、8KB ブロックの 1MB 読み取りでは、128 ブロックの読み取りが必要になるため、このパラメータのデフォルト値は 128 になります。

富士通がお客様環境で確認したデータベースパフォーマンスの問題のほとんどは、このパラメータの設定が誤っていることに起因していました。Oracle バージョン 8 および 9 でこの値を変更する正当な理由があったということです。その結果、データベースが Oracle 10 以降にアップグレードされたために、パラメータが `init.ora` ファイルに知らずに存在する可能性があります。レガシー設定の 8 または 16 は、デフォルト値の 128 と比較して、シーケンシャル I/O パフォーマンスに大きな影響を与えます。

### ■ ベストプラクティス

`init.ora` ファイルに `db_file_multiblock_read_count` パラメータを含めないことをお薦めします。このパラメータを変更することでパフォーマンスが向上したことはなく、むしろ、シーケンシャル I/O スループットに明らかな障害が発生する場合が多くあります。

# 5.3　リドゥブロックサイズ

Oracle では、512 バイトまたは 4KB のリドゥブロックサイズがサポートされています。デフォルトは 512 バイトです。このサイズによりリドゥ操作中に書き込まれるデータ量が最小化されるため、最適なオプションは 512 バイトになると予想されます。ただし、4KB のサイズでは、非常に高いロギングレートでパフォーマンスが向上する可能性があります。たとえば、リドゥのブロックサイズが大きいほど、リドゥロギングが 50MBps の単一データベースの方が効率的な場合があります。リドゥロギングの総量が大きいデータベースを多数サポートするストレージシステムでは、4KB のリドゥブロックサイズに利点があります。これは、4KB ブロックの一部のみを更新する必要がある場合に、この設定によって非効率的な部分 I/O 処理が排除されるためです。

すべての I/O 操作がリドゥログブロックサイズの単一ユニットで実行されるというわけではありません。非常に高いロギングレートでは、通常、データベースは複数のリドゥブロックで構成される非常に大きな I/O 操作を実行します。通常、これらのリドゥブロックの実際のサイズはロギングの効率に影響しません。

## ■ ベストプラクティス

富士通では、アプリケーションのドキュメント化された要件や、富士通または Oracle カスタマーサポートによる推奨事項などの理由から、デフォルトのブロックサイズの変更のみを推奨しています。

# 5.4　チェックサムとデータの整合性

一般に投げかけられる質問の 1 つは、データベースのデータ保全性をどのようにして保護するかということです。この質問は、Oracle RMAN ストリーミングバックアップの使用に慣れているお客様が Snapshot ベースのバックアップに移行する場合に特によくあります。特に、RMAN はバックアップ操作中に整合性チェックを実行します。この機能には価値がありますが、主な利点は、最新のストレージアレイで使用されていないデータベースの場合です。Oracle データベースに物理ディスクを使用している場合、ディスクが古くなると最終的に破損が発生することはほぼ確実です。この問題は、実際のストレージアレイではアレイベースのチェックサムによって解決されます。

実際のストレージアレイでは、複数のレベルでチェックサムを使用することでデータの整合性が保護されます。IP ベースのネットワークでデータが破損した場合、TCP 層はパケットデータを拒否し、再送信を要求します。FC プロトコルには、カプセル化された SCSI データと同様にチェックサムが含まれます。アレイ上に配置された後、SANtricity OS が RAID およびチェックサム保護を備えています。破損は発生する可能性がありますが、ほとんどのエンタープライズアレイと同様に、検出されて修正されます。通常、ドライブ全体に障害が発生すると RAID の再構築が要求されますが、データベースの整合性は影響を受けません。まれに、SANtricity OS がチェックサムエラーを検出することがあります。これは、ディスク上のデータが破損していることを意味します。その後、ディスクが故障し、RAID の再構築が開始されます。ここでも、データ整合性は影響を受けません。

Oracle データファイルおよびリドゥログのアーキテクチャは、極端な状況下でも最高レベルのデータ整合性を提供するように設計されています。最も基本的なレベルでは、Oracle ブロックには、ほとんどすべての I/O に対するチェックサムと基本的な論理チェックが含まれます。Oracle がクラッシュしていない場合またはテーブル領域をオフラインにしていない場合、データは変更されません。データ整合性チェックのレベルは調整可能であり、Oracle を構成して書き込みを確認することもできます。

その結果、ほとんどすべてのクラッシュおよび障害シナリオを回復でき、非常にまれな回復不能な状況では、破損が即座に検出されます。

Oracle データベースを使用している富士通のお客様のほとんどは、Snapshot ベースのバックアップへの移行後、RMAN などのバックアップ製品の使用を中止しています。SMO でブロックレベルのリカバリを実行するために RMAN を使用できるオプションはまだあります。ただし、日常的には、RMAN、NetBackup、およびその他の製品は、月単位または四半期単位のアーカイブコピーを作成するためにごくまれにしか使用されません。

一部のお客様は、`dbv` を定期的に実行して、既存のデータベースの整合性チェックを実施します。不要な I/O 負荷が発生するため、この方法は推奨しません。前述のように、データベースで以前に問題が発生していなかった場合、`dbv` が問題を検出する可能性は 0 に近く、このユーティリティを使用すると、ネットワークとストレージシステムに非常に大きなシーケンシャル I/O 負荷がかかります。Oracle の既知のバグにさらされるなど、破損が存在すると考えられる理由がない限り、`dbv` を実行する理由はありません。

# 第 6 章
# サイジング

Oracle のパフォーマンスは I/O に重点が置かれてきました。従来、ユーザーはスピンドルの数を増やすか、スピンドルの回転速度を上げることで、このパフォーマンスを向上させていました。ETERNUS AB series フラッシュアレイの登場により、SSD を使用することでパフォーマンスを向上させることができます。

## 6.1 ETERNUS AB series I/O の概要

ETERNUS AB series ストレージシステムの全体的なパフォーマンスに影響を与える要因には、ネットワークインフラストラクチャなどの物理コンポーネントや、基盤となるストレージ自体の構成などがあります。一般的に、ストレージシステムのパフォーマンス調整には、40/30/30 ルールが適用されます。調整および設定の 40% はストレージシステムレベル、30% はファイルシステムレベル、30% はアプリケーションレベルで行うというものです。次の項では、ストレージシステムの仕様に関連する 40% について説明します。ファイルシステムおよびアプリケーションレベルの場合、一般的な考慮事項には次のものが含まれます。

- **I/O サイズ**
  ETERNUS AB series ストレージシステムは、主に応答性に優れたシステムです。I/O 操作を完了するには、ホストがその操作を要求する必要があります。ホストからの各要求の I/O サイズは、IOPS 数またはスループット（メガバイト / 秒 [MBps] またはギガバイト / 秒 [GBps]）に大きな影響を与える可能性があります。一般的に、I/O が大きいほど、IOPS の数は少なくなり、MBps も大きくなります。その逆も同様です。この関係は、スループット =IOPS´I/O サイズという計算式で定義されます。
- **読み取り要求と書き込み要求**
  I/O サイズに加えて、ストレージシステムレベルで処理される読み取り I/O 要求と書き込み I/O 要求の割合もストレージシステムに影響を与える可能性があります。ソリューションを設計するときは、この割合も考慮する必要があります。
- **シーケンシャルデータストリームとランダムデータストリーム**
  基盤となるディスクメディアの論理ブロックアドレスへのホストリクエストは、シーケンシャルまたはランダムに行うことができ、ストレージシステムレベルでのパフォーマンスに大きな影響を与えます。このシーケンスは、物理メディアが最小限のレイテンシで要求に効果的に応答できるかどうかに影響します。また、このシーケンスは、ストレージシステムのキャッシュアルゴリズムの有効性にも影響します。例外として、ランダム要求のレイテンシが長くならないのは、レイテンシが機械的に発生しない SSD の場合です。
- **同時 I/O 処理数**
  特定のボリュームに適用される未処理の I/O 操作の数は、ファイルシステムが raw I/O、バッファ I/O、または直接 I/O を使用するかどうかなど、いくつかの要因によって異なります。一般に、ETERNUS AB series ストレージシステム内のほとんどのボリュームは、複数のドライブでストライピングされています。個々のディスクに最小限の未処理 I/O を割り当てた場合、ストレージシステム内のリソースが十分に活用されず、必要とされるパフォーマンスを確保できない場合があります。

ETERNUS AB series テクノロジーを初めて使用する場合は、RAID10、RAID5、RAID6、および DDP テクノロジーの違いを確認することをお勧めします。表 6.1 は、さまざまな RAID レベルの有効容量を比較したものです。完全を期すために、ETERNUS AB series システムでサポートされているすべての RAID レベルを示しています。

表 6.1　RAID レベルごとの有効容量の比較

| 求められる機能 | RAID 0 | RAID 1<br>RAID 10 | RAID 5 | RAID 6 | DDP |
|---|---|---|---|---|---|
| 有効容量 | 100% | 50% | (N-1)/N<br>(N はボリュームグループ内で選択されているドライブ数 ) | (N-2)/N<br>(N はボリュームグループ内で選択されているドライブ数 ) | 80% から、選択されている予約済み容量を引いた容量 |

# 6.2　I/O の見積もり

データベースのサイズを設定するときは、システムに必要な I/O 操作の数を見積もることが重要です。ここでは、データベースインスタンスのパフォーマンスを許容範囲内に維持する方法について説明します。システムの実際の物理 I/O 番号を取得できない場合は、I/O を見積もる必要があります。これは通常、新しいシステムを構築中に発生します。次の項では、I/O を見積もる計算式を示します。

## 6.2.1　新規の OLTP データベースシステム

次の項目を考慮する必要があります。

- ビジネストランザクション
- ビジネストランザクションの期間
- アプリケーションの許容可能なレイテンシ
- トランザクションとシステムトランザクション (I/O) のおおよその割合

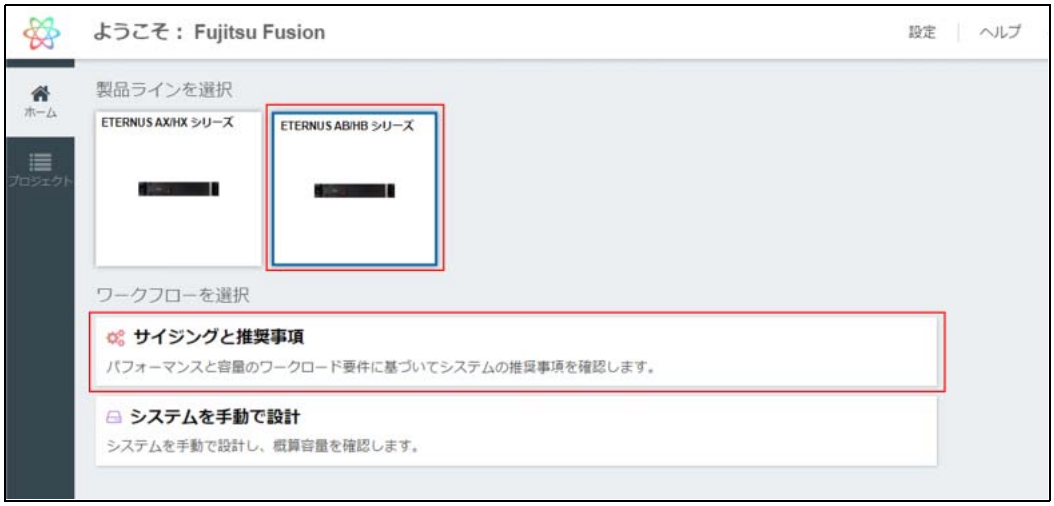
次の例は、ETERNUS AB series で Oracle のサイズを見積もる最も簡単な方法を示しています。

**手順 ▶▶▶**

1. ビジネストランザクションの数を見積もります。たとえば、1 日あたりのビジネストランザクション数を、600,000,000 と見積もります。ビジネスは、24 時間 365 日体制で運営されます。
2. 1 つのビジネストランザクションで 25 個の I/O 操作が作成されるとします。したがって、この例では、1 日あたりの I/O 操作は、600,000,000 x 25 = 15,000,000,000 I/O です。
3. 標準的なストレージサイジングツールを使用するには、IOPS 情報が必要です。この例では、(15,000,000,000) / 86,400 = 173,612 IOPS です。
4. アプリケーションの許容可能なレイテンシを考慮する必要があります。たとえば、アプリケーションは 1 ミリ秒未満のレイテンシを必要とする場合があります。

5 ETERNUS AB/HB series のプロダクトラインを選択した後、この数値を「サイジングと推奨事項」オプションで Fusion ツールに入力します。図 6.1 を参照してください。

図 6.1　Fusion を使用したシステムのサイズ変更と推奨




6 サイジングツールを使用するには、Read と Write の割合も見積もる必要があります。OLTP データベースでは、一般的に Read が 80%、Write が 20% です。

## 6.2.2　既存の OLTP データベースシステム

既存のデータベース環境のサイジングをする場合は、ワークロードのタイプを理解し、統計データを解釈すると役立ちます。システムに負荷がかかるピーク時に統計を収集することが重要です。Oracle AWR レポートには、システムを監視する時間枠のピークが表示されます。

システムの IOPS またはスループット (MBps) がキャプチャされたら、前述のように Fusion ツールに入力できます。

# 第 7 章

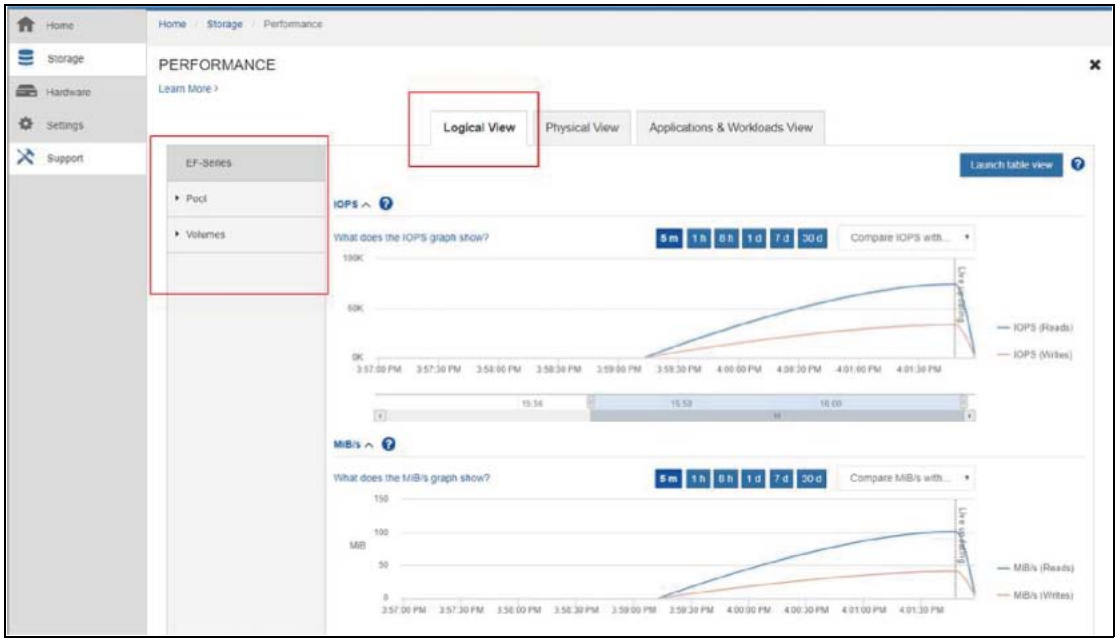
# SANtricity System Manager を使用した ETERNUS AB series のパフォーマンス監視

ストレージシステム操作中は、ストレージシステムのパフォーマンスを監視すると便利です。SANtricity System Manager を使用すると、ETERNUS AB series のパフォーマンスデータをテキスト形式とグラフィカル形式の両方のダッシュボード形式で表示できます。

SANtricity System Manager は、論理レベル、物理レベル、およびアプリケーションレベルのワークロードをキャプチャできる、優れたパフォーマンス監視機能を提供します。アプリケーションとワークロードのタグ付け、パフォーマンスデータの強化、組み込みモニタ、ボリューム使用率のグラフィカル表示などの新機能が追加されています。パフォーマンスモニタを表示するには、SANtricity System Manager で「View Performance Details」を選択します。3 つの主要なカテゴリのパフォーマンスを確認できます。

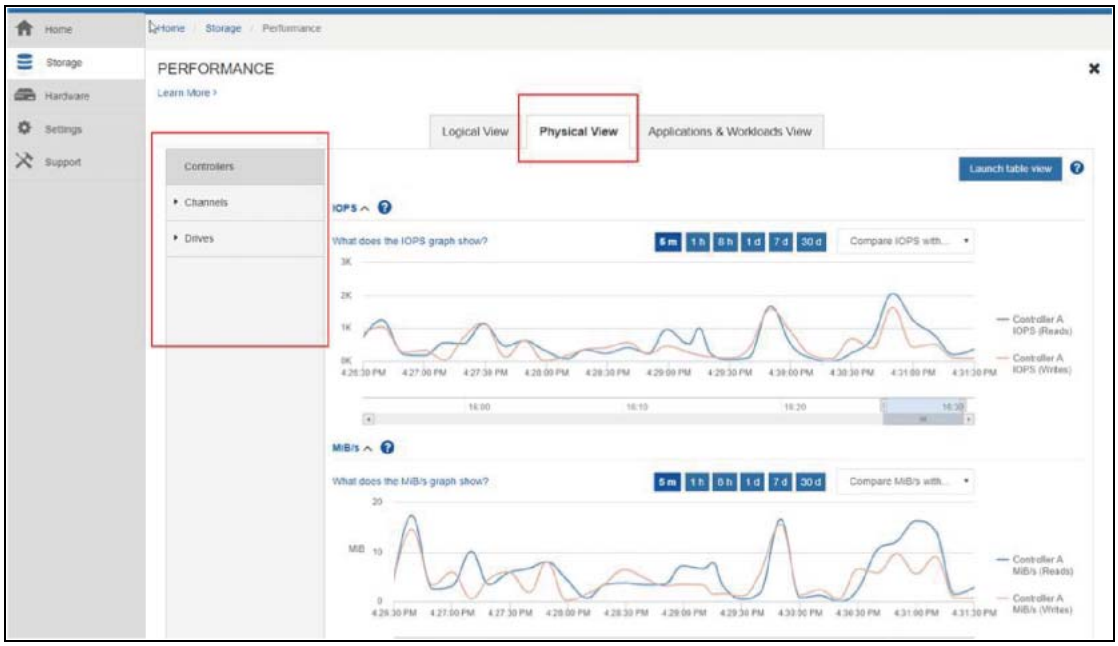
- 「Logical View」を使用すると、プール、ボリュームグループ、またはボリュームによって情報をフィルタリングできます。図 7.1 を参照してください。

図 7.1　Logical View を使用したパフォーマンスモニタの例



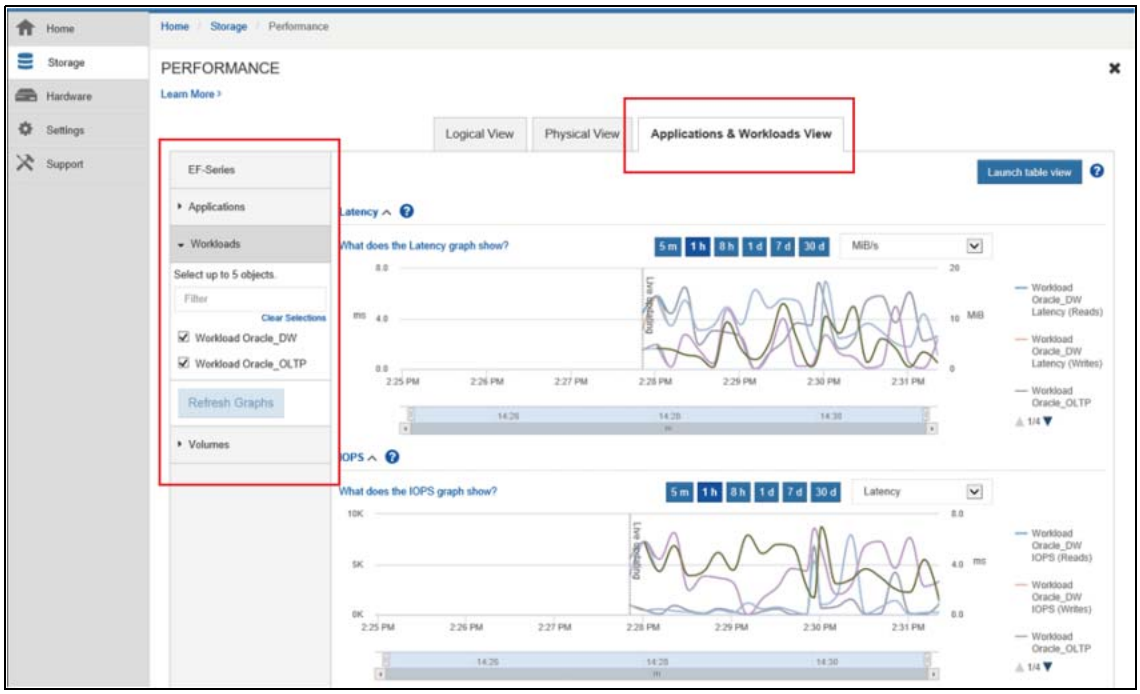
- 「Physical View」を使用すると、ホストチャネルとドライブによって情報をフィルタリングできます。図 7.2 を参照してください。

図 7.2　Physical View を使用したパフォーマンスモニタの例




- 「Applications & Workloads View」を使用すると、アプリケーション、ワークロード、およびワークロード内のボリュームによって情報をフィルタリングできます。図 7.3 を参照してください。

図 7.3　Applications & Workloads View を使用したパフォーマンスモニタの例



# 第 8 章
# 推奨事項

ETERNUS AB series では、パフォーマンスを最大にするためにデフォルト設定をほとんど変更する必要がありません。

## 8.1 ストレージ

表 8.1 に、ETERNUS AB series ストレージアレイの基本的なチューニングパラメータ、デフォルト設定、および推奨事項を示します。

表 8.1 ストレージチューニングパラメータ

| 名称 | コンテキスト | デフォルト | 使用可能なオプション | 推奨事項 |
|---|---|---|---|---|
| キャッシュブロックサイズ | アレイ | 32K | 4K、8K、16K、および 32K | 32K |
| 自動キャッシュフラッシュ | アレイ | 80% | 0% ～ 100% | 80% |
| セグメントサイズ | ボリューム | 128K | 32K、64K、128K、256K、512K | 64K または 128K |
| リードキャッシュ | ボリューム | 有効 | 有効 / 無効 | 有効 |
| 動的キャッシュ読み取りプリフェッチ | ボリューム | 有効 | 有効 / 無効 | 無効 |
| ライトキャッシュ | ボリューム | 有効 | 有効 / 無効 | 有効 |
| ライトキャッシュミラーリング | ボリューム | 有効 | 有効 / 無効 | 有効 |

## 8.2 Oracle ASM

表 8.2 に、Oracle ASM およびデータベースインスタンスの基本チューニングパラメータを示します。

表 8.2 Oracle ASM およびインスタンスの設定

| 名称 | デフォルト | 使用可能なオプション | 推奨事項 |
|---|---|---|---|
| ASM AU | 1MB | 1MB ～ 64MB | 64MB |
| ファイルシステム I/O オプション | なし | なし | 未使用 |
| db_file_multiblock_read_count | 128 | 8 ～ 128 | デフォルト |

# 8.3　Linux と NVMe-oF

表 8.3 は Linux の基本的なチューニングパラメータを示しています。Linux OS と NVMe-oF の完全なセットアップについては、富士通マニュアルサイトの「ETERNUS AB/HB series 構成エクスプレスガイド （Linux®）」を参照してください。

表 8.3　Linux チューニングパラメータ

| 名称 | デフォルト | 使用可能なオプション | 推奨事項 |
|---|---|---|---|
| I/O スケジューラ | cfq | cfq、anticipatory、deadline、または noop | noop |

# 第 9 章
# まとめ

本書では、高パフォーマンスの Oracle データベース向けの ETERNUS AB series ソリューションについて説明します。このソリューションでは、クラス最高レベルのエンドツーエンドの最新の SAN および NVMe テクノロジーを活用して、ビジネスクリティカルな IT サービスを今日提供するとともに、将来に備えます。

ETERNUS AB series により、富士通は将来のニーズにも対応し、現在も使用可能な SAN アレイを構築しました。これにより、現在の運用プロセスと手順に容易に実装できます。

AB3100 ストレージアレイに SAS 拡張機能を追加すると、現在および将来のニーズに対応する追加機能が提供されます。コントローラドライブトレイには、最大 240 台の大容量 NL-SAS ドライブまたは 96 台の SAS SSD ドライブと、24 台の NVMe ドライブを搭載できます。Oracle データベースの場合は、NVMe SSD ドライブで高パフォーマンスのデータベースを実行し、データを NL-SAS ドライブにバックアップできます。Oracle の DBA が希望する場合は、NL-SAS ドライブまたは SAS SSD ドライブを、Oracle パーティショニングを使用してデータベースのコールド層またはウォーム層としてデプロイできます。

ETERNUS AB series フラッシュアレイは、高パフォーマンス、一貫した低レイテンシ、および高度な HA 機能を提供して市場をリードする製品です。

アレイは、組み込みの System Manager を使用して容易にプロビジョニングできます。準備するシステムが多数ある場合は、REST API を使用してアレイをプロビジョニングすることもできます。

ETERNUS AB series フラッシュアレイには、堅牢な監視機能も組み込まれています。この機能を使用すると、パフォーマンスの問題を論理レベル、物理レベル、およびアプリケーションレベルでトラブルシューティングできます。

ETERNUS AB series フラッシュアレイは、極度のレイテンシ要件を解決するだけでなく、Oracle データベースに関する課題も次のように解決します。

- 既存のアプリケーションのパフォーマンスを大幅に向上させ、アプリケーションを再設計することなく IOPS あたりのコストを削減します。
- RAID10 と DDP により、Oracle のパフォーマンスが向上します。
- レスポンスタイムの短縮により、ユーザーの生産性が向上し、ビジネス効率が向上します。
- Oracle ASM は、高可用性およびディザスタリカバリのための代替ソリューションを提供します。

# 付録 A
# Oracle データベースバックアップ用の StorageGRID オブジェクトストレージ

データベースには、お客様情報、企業の財務情報、人事情報、その他の重要なデータソースなど、お客様の最も重要な資産が保存されています。今日における、常時接続されているグローバルなビジネス環境においては、災害からデータを保護するため、コストパフォーマンスと信頼性に優れた戦略が求められます。優れたデータ保護を実現するには、常時使用可能なビジネスクリティカルなバックアップを遠隔地に確保する必要があります。これは従来、バックアップをテープに書き込み、テープを遠隔地に保管することで実装されていました。

FUJITSU Hybrid IT Service for AWS の Amazon S3 や Azure Blob のようなクラウドオブジェクトストレージの全体的な価格と運用特性は、テープを遠隔地に輸送することに代わる魅力的な選択肢となっています。優れたクラウドインフラストラクチャは、ストレージの冗長性、セキュリティ、可用性、および拡張性を地理的に分散させて提供するため、可用性の損失を最小限に抑えながら、さまざまな有害事象を吸収できます。ただし、ネットワーク経由のクラウドバックアップに対する唯一の問題点は、ネットワーク帯域幅が制限されていると、大規模な本番 Oracle データベースのフルバックアップなど、大量のデータの転送速度が低下する可能性があることです。

これに代わって、Oracle のバックアップをオンサイトのオブジェクトであるストレージシステムに作成し、情報ライフライクル管理 (ILM) ポリシーに基づいて、別の遠隔地にあるクラウドプロバイダやデータセンターなどに、領域を越えてレプリケートする方法があります。これにより、アプリケーション開発の簡易的なプロビジョニングや、クラウドまたは遠隔データセンターでのテスト、および媒体故障やそのほかの災害からの保護を受けながら、全体の TCO を半分にすることが可能です。

備 考

Oracle Secure Backup Cloud Module には、特殊用途ライセンスが必要です。詳細は、Oracle® Secure Backup のライセンス情報を参照してください。

## A.1 当社のソリューション

StorageGRID は、これらの両方のニーズに対応できる業界トップのオブジェクトストレージソフトウェアソリューションです。この方法では、組み込みの暗号化や圧縮などの Oracle Recovery Manager (RMAN) のネイティブ機能を保持しながら、高い信頼性、継続的なアクセス性、無制限の拡張、遠隔地のストレージ費用の削減などのメリットを提供する、ハイブリッドクラウドソリューションが提供されます。

Amazon Web Services (AWS) は Oracle が提携する最初のクラウドベンダーで、FUJITSU Hybrid IT Service for AWS の Amazon S3 のオブジェクトストレージにデータベースバックアップを直接書き込みます。Oracle Secure Backup Cloud Module には、RMAN SBT インターフェースがバックアップを StorageGRID などの FUJITSU Hybrid IT Service for AWS の Amazon S3 または S3 互換のストレージに書き込むために必要なライブラリおよびウォレットキーを構成するのに役立つ、jar ファイルがあります。

# A.2 クラウドモジュールの構成と StorageGRID へのバックアップを書き込む 5 つのステップ

シンプルな 5 つのステップで Cloud Module 全体を構成し、バックアップを StorageGRID に書き込みます。

## 手順 ▶▶▶

1. Oracle サイトからパッケージをダウンロードし、解凍します。ダウンロードしたパッケージには、Oracle ユーザーが実行する jar ファイルが含まれています。
2. $ORACLE_HOME/dbs の配下に `osbws_wallet` ディレクトリを作成します。このディレクトリは、StorageGRID テナントアカウントのバケットへのアクセスに必要な、ウォレットキーの保存に使用されます。

```
drwxr-xr-x 2 oracle oinstall     4096 Mar  6 13:58 osbws_wallet
```

3. 次の必須パラメータを指定して jar ファイルを実行します。

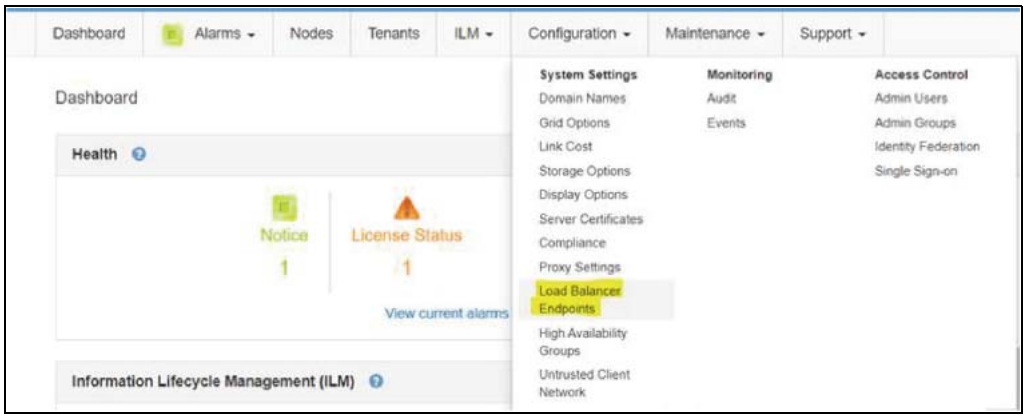
```
-bash-4.2$ java -jar osbws_install.jar -AWSID FXEBC3PXWZU1xxxx -AWSKey bZg2r0BqNGHZ9FbIeYcIx4rjYxxxxxxx -walletDir /orabin/product/18.0.0.0/dbhome_1/dbs/osbws_wallet -libDir
$ORACLE_HOME/lib -location us-east-1 -awsEndpoint webscaledemo.netapp.com  -useSigV2 -useHttps

Oracle Secure Backup Web Service Install Tool, build 12.2.0.1.0DBBKPCSBP_2018-06-12
AWS credentials are valid.
Oracle Secure Backup Web Service wallet created in directory /orabin/product/18.0.0.0/dbhome_1/dbs/osbws_wallet.
Oracle Secure Backup Web Service initialization file /orabin/product/18.0.0.0/dbhome_1/dbs/osbwsNTAP.ora created.
Downloading Oracle Secure Backup Web Service Software Library from file osbws_linux64.zip.
Download complete.
-bash-4.2$
```

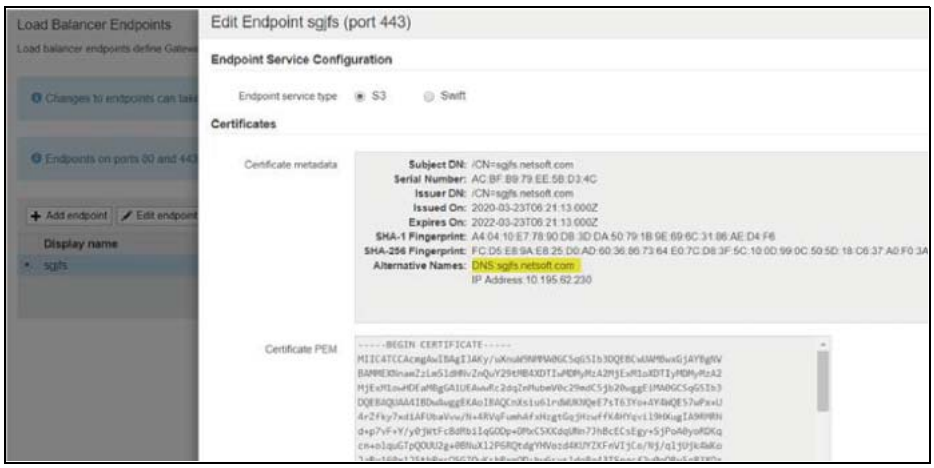
```
-bash-4.2$ java -jar osbws_install.jar -AWSID FXEBC3PXWZU1xxxx -AWSKey bZg2r0BqNGHZ9FbIeYcIx4rjYxxxxxxx -walletDir /orabin/product/18.0.0.0/dbhome_1/dbs/osbws_wallet -libDir
$ORACLE_HOME/lib -location us-east-1 -awsEndpoint webscaledemo.netapp.com  -useSigV2 -useHttps

Oracle Secure Backup Web Service Install Tool, build 12.2.0.1.0DBBKPCSBP_2018-06-12
AWS credentials are valid.
Oracle Secure Backup Web Service wallet created in directory /orabin/product/18.0.0.0/dbhome_1/dbs/osbws_wallet.
Oracle Secure Backup Web Service initialization file /orabin/product/18.0.0.0/dbhome_1/dbs/osbwsNTAP.ora created.
Downloading Oracle Secure Backup Web Service Software Library from file osbws_linux64.zip.
Download complete.
-bash-4.2$
```

AWSID および AWSKey パラメータは、StorageGRID の特定のテナントにアクセスするために必要な、アクセス ID およびアクセスキーを参照します。location パラメータは常に `us-east-1` にする必要があります。これは、Geo の種類に関係なく、StorageGRID に渡されるデフォルトの場所であるためです。他のクラウドオブジェクトストアでは、パラメータが異なる場合があります。awsEndpoint パラメータは、バケットへのアクセスに必要なロードバランサーのエンドポイント URL(HTTP または HTTPS) を表します。




複数の URL をホストし、DNS 内の適切なレコードと自己署名または商用証明書を使用して、バケット内のオブジェクトにアクセスすることもできます。




StorageGRID を使用すると、ロードバランサーのエンドポイントを作成するときに、独自の自己署名証明書を生成するか、商用証明書をアップロードすることができます。自己署名証明書を使用している場合は、Oracle ホストによって信頼されていることを確認する必要があります。それ以外の場合、jar ファイルの実行は失敗します。

次の例は、keytool を使用して自己署名証明書をインポートし、信頼されるようにする方法を示しています。

**3-1** StorageGRID によって生成された自己署名証明書を渡すには、-file パラメーターを使用します。

```
[root@jfs8 ~]# sudo keytool -import -trustcacerts -alias selfsigned -file /root/public.crt -keystore /usr/lib/jvm/java-1.8.0-openjdk-1.8.0.242.b08-0.el7_7.x86_64/jre/lib/security/cacerts
Enter keystore password:
Owner: CN=*.webscaledemo-rtp.netapp.com
Issuer: CN=*.webscaledemo-rtp.netapp.com
Serial number: fe5ae147d2480423
Valid from: Tue Mar 10 11:57:09 EDT 2020 until: Mon Mar 18 11:57:09 EDT 2030
Certificate fingerprints:
         MD5:  7A:14:53:F6:5D:1E:07:9B:4C:08:85:C1:5A:BF:95:31
         SHA1: 7C:EF:9B:D9:E3:18:4A:72:09:6A:B8:D4:BC:26:DA:29:88:4B:41:D0
         SHA256: 4E:BB:EE:CF:7B:23:EC:85:18:60:06:A9:BA:F1:7C:1F:D1:00:97:24:C3:6C:EF:9E:C6:46:A7:67:7E:B3:51:A0
Signature algorithm name: SHA256withRSA
Subject Public Key Algorithm: 2048-bit RSA key
Version: 3

Extensions:

#1: ObjectId: 2.5.29.17 Criticality=false
SubjectAlternativeName [
  DNSName: *.webscaledemo-rtp.netapp.com
  DNSName: webscaledemo-rtp.netapp.com
  IPAddress: 216.240.31.193
]

Trust this certificate? [no]:  yes
Certificate was added to keystore
[root@jfs8 ~]#
```

**4** jar の実行が成功したら、次に示すように、Oracle によって 4 つのファイルが作成されたことを確認します。
ここでは、$ORACLE_HOME/dbsの場所に3つのファイルがあり、そのうちの2つは`osbws_wallet`ディレクトリのキーに関連しています。もう一つは初期化ファイル`osbws<dbname>.ora`に関係しており、ホストの詳細とウォレットを記録しています。

```
drwxr-xr-x 2 oracle oinstall        4096 Feb 26 04:14 osbws_wallet
-rw-r--r-- 1 oracle oinstall         147 Feb 26 04:14 osbwsNTAP.ora
-bash-4.2$ pwd
/orabin/product/18.0.0.0/dbhome_1/dbs
-bash-4.2$ cat osbwsNTAP.ora
OSB_WS_HOST=http://s3.amazonaws.com
OSB_WS_WALLET='location=file:/orabin/product/18.0.0.0/dbhome_1/dbs/osbws_wallet CREDENTIAL_ALIAS=ng-hcl-r_aws'
-bash-4.2$ cd osbws_wallet
-bash-4.2$ ls
cwallet.sso  cwallet.sso.lck
```

4 番目のファイルは RMAN で使用されるライブラリで、$ORACLE_HOME/lib ディレクトリに`libosbws.so`という名前で生成されます。

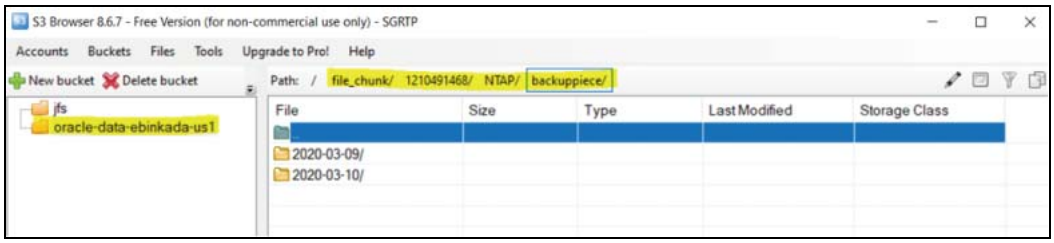
```
-rw-r--r-- 1 oracle oinstall  88218012 Feb 26 04:14 libosbws.so
-rw-r--r-- 1 oracle oinstall       286 Feb 26 04:14 metadata.xml
-bash-4.2$ pwd
/orabin/product/18.0.0.0/dbhome_1/lib
```

**5** 最後に、RMAN コマンドを実行して、Oracle バックアップを StorageGRID に直接送信するテストを行います。

```
RMAN> run {
            allocate channel ch1 type
            sbt parms='SBT_LIBRARY=/orabin/product/18.0.0.0/dbhome_1/lib/libosbws.so,
            SBT_PARMS=(OSB_WS_PFILE=/orabin/product/18.0.0.0/dbhome_1/dbs/osbwsNTAP.ora)';
            BACKUP AS BACKUPSET DATABASE PLUS ARCHIVELOG;
}
2> 3> 4> 5> 6>
using target database control file instead of recovery catalog
allocated channel: ch1
channel ch1: SID=301 device type=SBT_TAPE
channel ch1: Oracle Secure Backup Web Services Library VER=12.2.0.2


Starting backup at 10-MAR-20
current log archived
channel ch1: starting archived log backup set
channel ch1: specifying archived log(s) in backup set
input archived log thread=1 sequence=14 RECID=1 STAMP=1034698297
input archived log thread=1 sequence=15 RECID=2 STAMP=1034698298
input archived log thread=1 sequence=16 RECID=3 STAMP=1034698299
```

バケットは RMAN によって自動的に作成されます。S3 ブラウザまたはサードパーティ製ツールを使用して、StorageGRID テナントバケット内のオブジェクトをチェックできます。




バックアップ・ピース / バックアップ・セットのメタデータがカタログ化されているため、RMAN でバックアップの詳細を表示することもできます。バックアップが S3 または他の S3 互換ストレージにあるかどうかを確認するには、リストバックアップの`Media`パラメーターを慎重に検索します。

```
BS Key  Type LV Size       Device Type Elapsed Time Completion Time
------- ---- -- ---------- ----------- ------------ ---------------
5       Full    1.31G      SBT_TAPE    00:01:20     26-FEB-20
        BP Key: 5   Status: AVAILABLE  Compressed: NO  Tag: TAG20200226T044355
        Handle: 05upflkb_1_1   Media: s3.amazonaws.com/oracle-data-nghclroo-1
  List of Datafiles in backup set 5
  File LV Type Ckp SCN    Ckp Time  Abs Fuz SCN Sparse Name
  ---- -- ---- ---------- --------- ----------- ------ ----
  1       Full 2019428    26-FEB-20              NO    /oradata/NTAP/system01.dbf
  3       Full 2019428    26-FEB-20              NO    /oradata/NTAP/sysaux01.dbf
  4       Full 2019428    26-FEB-20              NO    /oradata/NTAP/undotbs01.dbf
  7       Full 2019428    26-FEB-20              NO    /oradata/NTAP/users01.dbf

BS Key  Type LV Size       Device Type Elapsed Time Completion Time
------- ---- -- ---------- ----------- ------------ ---------------
6       Full    10.25M     SBT_TAPE    00:00:01     26-FEB-20
        BP Key: 6   Status: AVAILABLE  Compressed: NO  Tag: TAG20200226T044521
        Handle: c-1210491468-20200226-02   Media: s3.amazonaws.com/oracle-data-nghclroo-1
  SPFILE Included: Modification time: 26-FEB-20
  SPFILE db_unique_name: NTAP
  Control File Included: Ckp SCN: 2019468      Ckp time: 26-FEB-20
```

# A.3 ベストプラクティス

以下のベストプラクティスは、バックアップまたはリストア ( あるいはその両方 ) 操作中の RMAN のパフォーマンスを向上させるのに役立ちます。

## ■ ベストプラクティス

- 複数の RMAN チャネルを使用して並列性を高め、ネットワークを完全に使用できるようにします。
- 圧縮により、ネットワーク帯域幅の制限を克服し、パフォーマンスを向上させることができます。
- 複数セクションのバックアップを使用します。Oracle Database では、セクションサイズパラメータを使用して、複数のチャネルで 1 つのファイルを同時にバックアップできます。
  BACKUP DEVICE TYPE SBT DATABASE SECTION SIZE 1g
- RMAN ブロックチェンジトラッキング機能を使用して、毎日の増分バックアップの性能を最適化します。
- オンプレミスデータベースの StorageGRID へのバックアップは、クラウドオブジェクトストレージに直接バックアップするよりもはるかに高速で効率的です。必要に応じて、ILM ポリシーに基づいてコールドオブジェクトをクラウドオブジェクトストアにレプリケートする場合は、クラウドミラー機能を使用できます。
- 本番データベースにアクティブな Oracle Data Guard インスタンスを使用している場合は、そのインスタンスを利用して増分 RMAN バックアップをより迅速に実行することができます。

# A.4 データベースのバックアップに StorageGRID を使用するメリット

## A.4.1 複数リージョンクラスタ、単一ネームスペース

StorageGRID は、単一のネームスペースを維持しつつ地理的に分散したデータセンターにまたがるクラスタを持つことができます。1 つのリージョンで取り込まれた Oracle データ・バックアップは他のリージョンに自動的にレプリケートされ、使用可能になります。1 つのリージョンが使用できなくなると、データは他のリージョンから自動的にクライアントとアプリケーションに提供されます。これは単一のネームスペースであるため、クライアントとアプリケーションはデータへのアクセス方法を変える必要がありません。

エンタープライズデータベースのお客様は、StorageGRID を両方の方法で使用できます。バックアップ (DR サイト ) の遠隔地コピーを保存することも、データを広く利用できるように複数のリージョンにまたがるワークフローを持つこともできます。

## A.4.2 データの決定性配置

各リージョンに必要なデータバックアップのレプリカの数を指定できます。企業のお客様は、プライマリデータセンターやその他のデータセンターでレプリカを指定できるように、ILM ポリシーを定義できます。

## A.4.3 消去コーディング

これは、データをフラグメントに分割し、冗長 ( パリティ ) フラグメントで符号化するデータ保護方法です。消去コーディングは、一般に、オブジェクトストレージシステム内のデータを保護するための、( レプリケーションと比較して ) よりスペース効率の高い方法です。StorageGRID 消去コーディングポリシーをマルチリージョンクラスタに適用することもできます。したがって、1 つ以上のノードまたは複数のドライブが失われた場合でも、他のリージョンからフラグメントを取得せずにオブジェクトを提供できます。

FUJITSU Storage
ETERNUS AB series オールフラッシュアレイ
ETERNUS AB series 上の Oracle Database

P3AG-6602-01Z0

発行年月　2022 年 5 月
発行責任　富士通株式会社

● 本書の内容は、改善のため事前連絡なしに変更することがあります。
● 本書の内容は、細心の注意を払って制作致しましたが、本書中の誤字、情報の抜け、本書情報の使用に起因する運用結果に関しましては、責任を負いかねますので予めご了承願います。
● 本書に記載されたデータの使用に起因する第三者の特許権およびその他の権利の侵害については、当社はその責を負いません。
● 無断転載を禁じます。

FUJITSU